

CC
CC -1- ENZYME REGULATION: Inhibited by DEFA (DEFF45).
CC -1- SUBUNIT: HETERODIMER OF A 45 kDa AND A 40 kDa SUBUNIT.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1- ALTERNATIVE PRODUCTS: 4 isoforms; Alpha (shown here), Beta, Gamma
CC and Delta; are produced by alternative splicing.
CC -1- SIMILARITY: CONTAINS 1 CIDE-N DOMAIN.
CC
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR	EMBL	AF064019	AAC39920.1	-
DR	EMBL	AB013918	BAA32250.1	-
DR	EMBL	AF039210	AAC39709.1	-
DR	EMBL	AB028911	BAB80447.1	-
DR	EMBL	AB028912	BAB80448.1	-
DR	EMBL	AB028913	BAB80449.1	-
HSSP	000273	11BX		

Genew;HGNC:2113; DFFB.
DR
MIM: 601883; -
DR InterPro: IPR003508; CAD.
DR Pfam: PF02017; CIDE-N; 1.
DR SMART: SM00266; CAD; 1.
DR Hydrolase: Nuclease; Nuclear protein; Apoptosis; Alternative splicing
KJ119696.1

[illegible]

SEQUENCE	336 AM	32103 MM	UB02F30D51209383	CR004
Query Match	100.0%	Score 1789	DB 1	length 338
Best Local Similarity	100.0%	Pred. NO. 5.3e-147		
Matches 338	Conservative 0	Mismatches 0	Indels 0	Gaps 0

09 .

1 MLOPKSVKTLALRSPRRFVGAGRSCEVLRKRGCLFOLPGRSRLCYEDTETLDEDF 60

1 MLQAPKSVKLALNLSPRRFVGAGRSCEVLRKRGCLFOLPGRSRLCYEDTETLDEDF 60

61 PSYDMMETVLITATCGAAGCACTGCTTTTGTCTCTATTGGTGCTTAACTTGGTACCTCATTA 100

09Y 121 DLININSONIAAETRAEDPPEFEGHESFQASKGLRSCSRISYLRREVSSYSTNGA 180
D6 61 PSVPDMNELVLTLTGQAMOGFVSDIRRLSFAFHEBOYGLIDAAOOLCDEQAPQBCORLTA 120

Db 121 DLHNSQNI~~AAETRAED~~PPWFEGLE~~SR~~FSKSGYL~~IR~~SCSE~~RL~~RSYL~~RE~~VSSYSTVG~~A~~ 180

Qy 181 EAQEEFLRV~~L~~SGSMCQRL~~NS~~SMQY~~NS~~YFDPGAKGGSRLCTPG~~WF~~SCQGPFDMD~~S~~CLSRHS 240

Db 181 EAQEEFLRV~~L~~SGSMCQRL~~NS~~SMQY~~NS~~YFDPGAKGGSRLCTPG~~WF~~SCQGPFDMD~~S~~CLSRHS 240

[illegible]

	QY
	301 LVHIVCHKKTTHLNCDEPSRIYYKPQTRLKKRPKRPO 338
D6	301 LVHIVCHKKTTHLNCDEPSRIYYKPQTRLKKRPKRPO 338

RESULT 2	DFEB_MOUSE	DFEB_MOUSE	STANDARD;	PRT;	344 AA.
ID	ID				
AC	054786;				
DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	DNA fragmentation factor 40 kDa subunit (EC 3.1.11.1) (DFE-40)				
DE	(Caspase-activated deoxyribonuclease) (Caspase-activated Dnase) (CAD)				
GN	DFEB OR CAD.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
NCBI	taxid=10090;				
TX					

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9808269; PubMed=9422506;
 RA EDLARI M., Sakahira H., Yokoyama H., Okawa K., Iwamatsu A., Nagata S.;
 RT "A caspase-activated DNase that degrades DNA during apoptosis, and its
 RT inhibitor ICAD.";
 RL Nature 391:43-50(1998).
 RN [2]

6/105

RP CHARACTERIZATION.
RX MEDLINE=98082976; PubMed=9422513;
RA Sakahira H., Enari M., Nagata S.;
RT "Role of ICAD cleavage in CAD activation and DNA degradation.";
RT Nature 391:96-99(1998).
CC -!- FUNCTION: Nuclease that induces DNA fragmentation and chromatin

CC condensation during apoptosis. Degrades naked DNA and induces
CC apoptotic morphology.
CC -1 ENZYME REGULATION: Inhibited by DFPA (DFP45).
CC -1 SUBUNIT: HETERODIMER OF A 45 kDa AND A 40 kDa SUBUNIT.
CC -1 SUBCELLULAR LOCATION: Cytoplasmic and nuclear.
CC -1 SIMILARITY: CONTAINS 1 CIDE-N DOMAIN.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isdb-sib.ch/announce/>
CC or send an email to license@isdb-sib.ch).

```
CC -----
DR EMBL; AB009377; BAA24977.1; -.
DR HSSP; 000273; 1IBX.
DR MGD; MGI:1196287; Dfcb.
DR InterPro; IPR003508; CAD.
```

DR Pfam: PF02017; ClIDE-N: 1.
DR SMART: SM00266; CAD: 1.
KM Hydrolase; Nuclease; Nuclear protein; Apoptosis.
PT DOMAIN 7 83 ClIDE-N: 1.
SQ SEQUENCE 344 AA; 39449 MW; E854B4E13EA139DE1 CRC64;

Query Match	79.2%	Score 1416;	DB 1;	Length 344;
Best Local Similarity	77.2%;	Pred. No. 8.2e-115;		
Matches	261;	Conservative	38;	Mismatches 2;
				Indels 1;

QY 61 PYPVDNAELVLLTGGAMOGTVSDIRRTSAFHEPVGGLIAAQQLLCDEQAQRRLTA 120
.....
Dd 4 VLRQPCKVKTRALHSACKFGVAARSCQELLRKCCVFQPLPMPSRLCLXEDGETVDDCF 63
.....
QY

Db 64 PGLPNDPAELLILLIAGETWNGTYSVDITRFRLSVENEPHAGVIAQAKKOLLSDQALROKLLA 123

```

181 EA0EEFRLVLSMCQRRLSMQYNGSYFDKRGAGGSRICPTPCWVFCOGPFDMDSCLRHS 240
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```
QY 241 INPYNRESILFSTWNLDDHIEKRTIPLVAIEKODGREVMEYGLTSENK 300
DB 244 INPYNRESILFSTWNLDDHIEKRTIPLVAIEKODGREVMEYGLTSENK 301
QY 301 LVHIVCHKRTKRLNCDPSRIYKPTRLKRPQVRK 338
DB 302 LVHIVCHKRTKRLNCDPSRIYKPTRLKRPQVRK 339

RESULT 3
DEFB_RAT STANDARD: PRT: 349 AA.
ID DEFB_RAT 099N34;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DNA fragmentation factor 40 kDa subunit (EC 3.-.-.-) (DEF-40)
(Caspase-activated deoxyribonuclease) (Caspase-activated DNase) (CAD).
DEFB OR CAD.
Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=21319171; PubMed=11425895;
RA Cao G., Pei W., Fan J., Stebler R.A., Luo Y., Nagayama T.,
Graham S.H., Yin X.M., Simon R.P., Chen J.,
*Caspase-activated DNase/DNA fragmentation factor 40 mediates
apoptotic DNA fragmentation in transient cerebral ischemia and in
neural cultures. J. Neurosci. 21:4678-4690(2001).
RL J. Neurosci. 21:4678-4690(2001).
CC -1- FUNCTION: Nuclease that induces DNA fragmentation and chromatin
condensation during apoptosis. Degrades naked DNA and induces
apoptotic morphology.
CC -1- ENZYME REGULATION: Inhibited by DEFA (DEF45).
CC -1- SUBUNIT: HETERODIMER OF A 45 kDa AND A 40 kDa SUBUNIT.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (by similarity).
CC -1- SIMILARITY: CONTAINS 1 CIDE-N DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sdb.ch/announce/
or send an email to license@sdb-sdb.ch).
CC -----
EMBL: AF136598; AAK16646.1; -
HSSP: Q00273; 11BX
DR InterPro: IPR003508; CAD.
DR Pfam: PF02017; CIDE-N.1.
DR SMART: SMO0266; CAD; 1.
DR KJ0417; Nuclease; Nuclear protein; Apoptosis.
SQ SEQUENCE 349 AA; 40096 MW; F125972A110E398 CRC64;

Query Match 77.8%; Score 1392; DB 1; Length 349;
Best Local Similarity 76.3%; Pred. No. 9.9e-113;
Matches 257; Conservative 38; Mismatches 40; Indels 2; Gaps 1;

QY 1 MLQPKSVKRLALSPKRFVAGRSQCEVLRKGLRQELRPERGSRRLCYEDTETEDYF 60
DB 4 VLQPKSVKRLALSPKRFVAGRSQCEVLRKGLRQELRPERGSRRLCYEDTETEDYF 63
QY 61 PSVDNAELVLTLLGOAMOGYSDIRRLLSAFHBPQVGLIOAQQLLCDEQAPOKORLLA 120
DB 64 PSVDNAELVLTLLGOAMOGYSDIRRLLSAFHBPQVGLIOAQQLLCDEQAPOKORLLA 123
QY 121 DLHIVSONIAETREDDPWFEGLESFRQSKSYLRYSCESIRSYLREVSYSYTYGA 180
DB 124 DLHIVSONIAETREDDPWFEGLESFRQSKSYLRYSCESIRSYLREVSYSYTYGA 183
```

```
QY 181 EAQEEFLRYLGSQCMALRSQNGSYPRGAKGSRICLTPGPFSCQGPFDMSLSHS 240
DB 184 AAREELRYLSSKCHLKVQYNGSTFPRGABASSRLCTEGNFSQGPFDMSLSHS 243
QY 241 INPYNRESILFSTWNLDDHIEKRTIPLVAIEKODGREVMEYGLTSENK 300
DB 244 INPYNRESILFSTWNLDDHIEKRTIPLVAIEKODGREVMEYGLTSENK 301
QY 301 LVHIVCHKRTKRLNCDPSRIYKPTRLKRPQVRK 337
DB 302 LVHIVCHKRTKRLNCDPSRIYKPTRLKRPQVRK 338

RESULT 4
PCNT_MOUSE STANDARD: PRT: 1920 AA.
ID PCNT_MOUSE 48725;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Pericentrin.
GN PCNT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94170365; PubMed=8124707;
RA Dorey S.J., Stein P., Evans L., Calarco P.D., Kirschner M.,
"Pericentrin, a highly conserved centrosome protein involved in
microtubule organization."
RL Cell 76:639-650(1994).
CC -1- FUNCTION: INTERAL COMPONENT OF THE FILAMENTOUS MATRIX OF THE
CENTROSOME INVOLVED IN THE INITIAL ESTABLISHMENT OF ORGANIZED
MICROTUBULE ARRAYS IN BOTH MITOSIS AND MEIOSIS.
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN KIDNEY, THYMUS AND
LIVER. LOW LEVELS IN BRAIN, MUSCLE, LUNG AND HEART.
CC -1- DOMAIN: COMPOSED OF A COILED-COIL CENTRAL REGION FLANKED BY NON-
HELICAL N- AND C-TERMINALS.
CC -1- SIMILARITY: STRONG, TO HUMAN KENRIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sdb.ch/announce/
or send an email to license@sdb-sdb.ch).
CC -----
EMBL: U05823; AAL17886.1; -
DR MGD: MGI:102722; Pcnt.
DR KJ0417; Nuclease; Nuclear protein; Apoptosis.
SQ SEQUENCE 1920 AA; 218337 MW; CF1DADEC5B73309 CRC64;

Query Match 6.1%; Score 108.5; DB 1; Length 1920;
Best Local Similarity 23.9%; Pred. No. 0.48;
Matches 56; Conservative 45; Mismatches 90; Indels 43; Gaps 11;

QY 2 LQKPKSVKRLALSPKRFVAGRSQCEVLRKGLRQELRPERGSRRLCYEDTETEDYF 51
DB 142 LQKPKSVKRLALSPKRFVAGRSQCEVLRKGLRQELRPERGSRRLCYEDTETEDYF 54
QY 52 GTEETPEYPSVDNAELVLTLLGOAMOGYSDIRRLLSAFHBPQVGLIOAQQLLCDEQA 107
DB 54 GTEETPEYPSVDNAELVLTLLGOAMOGYSDIRRLLSAFHBPQVGLIOAQQLLCDEQA 110
QY 198 AAEKELRSEMEKNQ-TIETLKQDWSERELCLENRELSTKQSEMEGLQSQPQKE 256
DB 198 AAEKELRSEMEKNQ-TIETLKQDWSERELCLENRELSTKQSEMEGLQSQPQKE 256
QY 108 CDEQAPOKORLL-----ADLHIVSONIAA-----ETRAEDPWFEGLESFRQSK 154
DB 108 CDEQAPOKORLL-----ADLHIVSONIAA-----ETRAEDPWFEGLESFRQSK 154
```

[illegible]

CC	various non-epithelial cells, especially mesenchymal cells.
CC	-1- SUBUNIT: HOMOPOLYMER.
CC	-1- PIR: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
CC	MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
CC	DIVISION, AT WHICH TIME VIMENTIN FILAMENTIS ARE SIGNIFICANTLY
CC	REORGANIZED.
CC	-1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-stb.ch/announce/ ,
CC	or send an email to license@isb-stb.ch).
CC	-----
DR	EMBL; M24849; AAA40555.1; -
DR	EMBL; X56397; CAA39607.1; -
DR	EMBL; M26251; AAA40556.1; -
DR	EMBL; 422526; CAA80251.1; -
DR	EMBL; X51438; CAA35803.1; -
DR	EMBL; Y07738; CAA69019.1; -
DR	EMBL; D50805; BAA19834.1; -
DR	PIR; J50241; J50241.
DR	PIR; S12774; S12774.
DR	SWISS-2DPAGE; P20152; MOUSE.
DR	MGI; MGI:98932; Vlm.
DR	InterPro: IPR001664; IF.
DR	Pfam: PF00038; filament; 1.
DR	PROSITE; PS00226; IF; 1.
KM	Intermediate filament; Coiled coil; Phosphorylation.
FT	INIT_MET 0
FT	DOMAIN 1 94 HEAD.
FT	DOMAIN 95 406 ROD.
FT	DOMAIN 407 465 TAIL.
FT	DOMAIN 95 130 COIL 1A.
FT	DOMAIN 131 152 LINKER 1.
FT	DOMAIN 153 244 COIL 1B.
FT	DOMAIN 245 267 LINKER 12.
FT	DOMAIN 268 406 COIL 2.
FT	MOD_RES 38 38 PHOSPHORYLATION (BY CAMK2).
FT	MOD_RES 82 82 PHOSPHORYLATION (BY CAMK2).
FT	CONFLICT 69 69 L-> S (IN REF. 6).
FT	CONFLICT 109 114 LNDRA-> ILIAEL (IN REF. 6).
FT	CONFLICT 155 156 EL-> DV (IN REF. 4).
FT	CONFLICT 163 163 L-> F (IN REF. 2).
FT	CONFLICT 337 337 E-> V (IN REF. 4).
FT	CONFLICT 373 373 E-> D (IN REF. 1).
SO	SEQUENCE 465 AA; 53556 MW; 9CD5CE4B39AD9632 CRC64;
Query Match	5.5%; Score 99; DB 1; Length 465;
Best Local Similarity	20.1%; Pred. No. 0.51;
Matches	63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;
QY	27 QEVLRKKGRLQQLPERGSRLLCYEDGTALTEDYFSPVDNAELVLTTLGQAMQGYVSDIR 86
DB	179 EDIMR--LRKRLQGE-----MIGREARSTYLSQFQVDVNMSTLADDERKVESIQEEL- 230
QY	87 RFLSAFHEPQVGLIOA-----QQLCDEQAPQORRLIADLLIHVNSONIAETRAEDPPV 141
DB	231 AFLKLHDEEIOELQAOIOEHOVQIDVVSFKPDLTALRLDVQOY-ESVAAKNIQEAEEW 289
QY	142 F-----EGLESRFQSKGYLR-----SCSRIR-----S 166
DB	290 YKSKFADLSEANRRNDALRQAKQESNERKROYOSLCEVDALKGTNESLEROMKMEEN 349
QY	167 YLREYSSYPSTVGAENAEFLRVLSMCOURLRSMQYNGSYEDRGAKGSGRLCTPEGWFS 226
DB	350 FALAEANVQDITIG-RLDDEIQNMKEEMARHLREYQ-----DLNV 388
QY	227 QGPPDMQSLSRHSINPISNRSR-----LFTWNLDHT-----IEKKTIIFTL 272
DB	389 KMAIDIEIATRKLL-----EGEESRSLTLPFFSSNLTAEINLESILPVDTSKRKL--L 442

OY 273 VEAIRKEDGREN 285
DB 443 IKVETRDGQVIN 455

RESULT 6

VIME CRIGR STANDARD; PRT; 448 AA.

AC P48670;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vimentin (Fragment).

GN VIM. Crictellus griseus (Chinese hamster).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Cricetulus.
NCBI_TaxID=10029;

RA MEDLINE=8329727; PubMed=6688458;
RA Bioemendal H., Quax W., Quax-Jeukens Y., Dodemont H., Ramaekers F.,
Dunja I., Benedetti L.;
RT "Organization and expression of the vimentin gene.";
RL Mol. Biol. Rep. 9:115-118(1983).

RP SEQUENCE OF 69-448 FROM N.A.
RH Rhodjakov A.L., Koonce M.P.;
RL Submitted (MAY-1995) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Vimentins are class-III intermediate filaments found in
CC various non-epithelial cells, especially mesenchymal cells.
CC -1- SUBUNIT: HOMOPOLYMER.

CC -1- PTM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
CC REORGANIZED (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M16718; AAA37029.1;
EMBL; X87227; CAA60679.1;
InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Phosphorylation.

FT NON_TER 1
FT DOMAIN 1
FT DOMAIN 78
FT DOMAIN 78
FT DOMAIN 390
FT DOMAIN 78
FT DOMAIN 113
FT DOMAIN 114
FT DOMAIN 135
FT DOMAIN 227
FT DOMAIN 250
FT DOMAIN 251
FT DOMAIN 389
FT SEQUENCE 448 AA; 51848 MM; 2518FDEF59B3B87 CRC64;

Query Match 5.53; Score 98; DB 1; Length 448;
Best Local Similarity 20.18; Pred NO. 0.59; Indels 90; Gaps 14;
Matches 63; Conservative 47; Mismatches 113;

OY 27 QEVLRKGLRFPOLPERGSRKLEDEGTLEDEYPPSPVDNAEVLTLTGAMOGGYSDIR 86
DB 162 EDITR---LREKLOEE---MLOREASTLOSFRQDVNASTLARLDLREKVSLOEET- 213

OY 87 RFLSAFEPQVGLIQA-----QQLLCDEQAPQORRLADLLHNVSONIAETRAEDPPW 141
DB 214 AFLKRLDEEIOELQAOIOEHQVQIDVSKDLRALRDVROQY-ESVAKRLQAEAEW 272
OY 142 F-----EGLESFRQSKGYLR-----SCESRIR-----S 166
DB 273 YKSKFADLSEANNNNDALRQAQESNEFRQVQSLCEVDALKTGNETLERQMRMEEN 332
OY 167 YLRSSYPTVGALEAOEELRYLSMOCORLRSMQYNSYFDRKAGKSLCPDEGFWSC 226
DB 333 FALEANQDTIG-RLADEIQNKKEKMAHREY-----DLNV 371
OY 227 QGPRMDSCLSRHSINYSNRRL-----FTWNLDT-----TEKRTITPL 272
DB 372 KALDIEATYRKIL---EGESRSLPLPNSLNLKNETNESLPIVTHSKRL--L 425
OY 273 VEAIRKEDGREN 285
DB 426 IKVETRDGQVIN 438

RESULT 7

VIME MESAU STANDARD; PRT; 464 AA.

AC P02544;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vimentin.

GN Mesocricetus auratus (Golden hamster).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
NCBI_TaxID=10036;

RA MEDLINE=84026520; PubMed=6194898;
RA Quax W.J., Egberts W.V., Hendriks W., Quax-Jeukens Y.E.F.M.,
Bioemendal H.;
RT "The structure of the vimentin gene.";
RL Cell 35:215-223(1983).

RP SEQUENCE FROM N.A.
RH TISSUE=Lens;
RL MEDLINE=83221633; PubMed=6304716;
RA Quax-Jeukens Y.E.F.M., Quax W.J., Bioemendal H.;
RT "Primary and secondary structure of hamster vimentin predicted from
RT the nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:3548-3552(1983).

CC -1- FUNCTION: Vimentins are class-III intermediate filaments found in
CC various non-epithelial cells, especially mesenchymal cells.
CC -1- SUBUNIT: HOMOPOLYMER.
CC -1- PTM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
CC REORGANIZED.

CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; K00927; AAA37104.1; JOINED.
EMBL; K00921; AAA37104.1; JOINED.
EMBL; K00922; AAA37104.1; JOINED.
EMBL; K00923; AAA37104.1; JOINED.
DR EMBL; K00924; AAA37104.1; JOINED.

DR EMBL; K00925; AAA37104.1; JOINED.
 DR EMBL; K00926; AAA37104.1; JOINED.
 DR PIR; A02959; VEHY.
 DR InterPro; IPR001664; IF.
 DR Pfam; PF00038; filament. 1.
 DR PROSITE; PS00226; IF. 1.
 DR Intermediate filament; Coiled coil; Acetylation; Phosphorylation.
 FT INT_MET 0 0
 FT DOMAIN 1 93 HEAD.
 FT DOMAIN 94 405 ROD.
 FT DOMAIN 406 464 TAIL.
 FT DOMAIN 94 129 COIL 1A.
 FT DOMAIN 130 151 LINKER 1.
 FT DOMAIN 152 243 COIL 1B.
 FT DOMAIN 244 266 LINKER 12.
 FT DOMAIN 267 405 COIL 2.
 FT MOD_RES 1 1 ACETYLATION.
 FT MOD_RES 38 38 PHOSPHORYLATION (BY CAM2) (BY
 MOD_RES 81 81 SIMILARITY).
 MOD_RES PHOSPHORYLATION (BY CAM2) (BY
 CONFLICT 42 42 L -> A (IN REF. 2).
 CONFLICT 115 115 Y -> D (IN REF. 2).
 CONFLICT 182 182 R -> I (IN REF. 2).
 FT SEQUENCE 464 AA; 53598 MW; 7AC417008C8D4776 CRC64;
 Query Match 5.5%; Score 98; DB 1; Length 464;
 Best Local Similarity 20.1%; Pred. No. 0.62;
 Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;
 QY 27 QEVLRKGLRQLEPGRSRLCLYEDGTELTEDYPSVNDALVLTGQAMQGVSDIR 86
 DB 178 EDIMR---LRKRLQEE---MLQREASTIQSFROVDNASTLRDLDERVESLQERI 229
 QY 87 RFLSAFHEPQVGLIQAA-----QQLCDEQAPQORRLADILHNVSONIAETRAEDPPW 141
 DB 230 AFLKKHDEEIOEQAOIQEOHQVIDVSKPDIALRDVROQY-ESVAANKIQEAEW 288
 QY 142 F-----EGLESRFQSKGYLRY-----SCESRIR---S 166
 DB 289 YKSKFADISEANNNALRQAKQESNEYRQVSLCEVDALKTGNESELRQREMEEN 348
 QY 167 YLRVSSPSTVGAQAEFLRVLSGMCQRLRSQYNSYDRAGKSGRLCTPEGWFS 226
 DB 349 FALPAAHQDTIG-RLQDEIQNMKEEAKRLREYO-----DILNV 387
 QY 227 QGPPDMSCLSRHSINPYSNRESRL-----FSTWNLDIH-----IEKKRTIIPTL 272
 DB 388 KMAIDIEIATYRKLL-----EGESRISLPLPFSLSLRETNLJESLPLVDTHSKRTL---L 441
 QY 273 VEAIKEDGQREV 285
 DB 442 IKVETRDGQVIN 454
 RESULT 8
 VIME_RAT STANDARD; PRT; 465 AA.
 AC P31000;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Vimentin.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fischer;
 RX MEDLINE=92171936; PubMed=1540169;
 RA Bussemakers M.J.G., Verhaegh G.W.C.T., van Bokhoven A.,

RA Debruyne F.M.J., Schalken J.A.;
 RT "Differential expression of vimentin in rat prostatic tumors";
 RL Biochem. Biophys. Res. Commun. 182:1254-1259(1992).
 RN (2)
 RP SEQUENCE OF 85-159 FROM N.A.
 RA Paine M.L.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Vimentins are class-III intermediate filaments found in
 CC various non-epithelial cells, especially mesenchymal cells.
 CC -1- SUBUNIT: HOMOPOLYMER.
 CC -1- PTM: ONE OF THE MOST PROMINENT PHOSPHORYLATIONS IN VARIOUS CELLS OF
 CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
 CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
 CC REORGANIZED.
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X62952; CAA44722.1; -;
 DR EMBL; M84481; AAA42339.1; -;
 DR PIR; S22119; S22119.
 DR PIR; J01389; J01389.
 DR InterPro; IPR001664; IF.
 DR Pfam; PF00038; filament. 1.
 DR PROSITE; PS00226; IF. 1.
 DR Intermediate filament; Coiled coil; Phosphorylation.
 DR KW Intermediate filament; BY SIMILARITY.
 FT INT_MET 0 0
 FT DOMAIN 1 94 HEAD.
 FT DOMAIN 95 406 ROD.
 FT DOMAIN 407 465 TAIL.
 FT DOMAIN 95 130 COIL 1A.
 FT DOMAIN 131 152 LINKER 1.
 FT DOMAIN 153 244 COIL 1B.
 FT DOMAIN 245 267 LINKER 12.
 FT DOMAIN 268 406 COIL 2.
 FT MOD_RES 38 38 PHOSPHORYLATION (BY CAM2) (BY
 FT MOD_RES 82 82 SIMILARITY).
 FT MOD_RES PHOSPHORYLATION (BY CAM2) (BY
 FT SIMILARITY).
 FT SEQUENCE 465 AA; 53601 MW; 6587FE7652CB3682 CRC64;
 Query Match 5.5%; Score 98; DB 1; Length 465;
 Best Local Similarity 20.1%; Pred. No. 0.62;
 Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;
 QY 27 QEVLRKGLRQLEPGRSRLCLYEDGTELTEDYPSVNDALVLTGQAMQGVSDIR 86
 DB 179 EDIMR---LRKRLQEE---MLQREASTIQSFROVDNASTLRDLDERVESLQERI 230
 QY 87 RFLSAFHEPQVGLIQAA-----QQLCDEQAPQORRLADILHNVSONIAETRAEDPPW 141
 DB 231 AFLKKHDEEIOEQAOIQEOHQVIDVSKPDIALRDVROQY-ESVAANKIQEAEW 289
 QY 142 F-----EGLESRFQSKGYLRY-----SCESRIR---S 166
 DB 290 YKSKFADISEANNNALRQAKQESNEYRQVSLCEVDALKTGNESELRQREMEEN 349
 QY 167 YLRVSSPSTVGAQAEFLRVLSGMCQRLRSQYNSYDRAGKSGRLCTPEGWFS 226
 DB 350 FALPAAHQDTIG-RLQDEIQNMKEEAKRLREYO-----DILNV 388
 QY 227 QGPPDMSCLSRHSINPYSNRESRL-----FSTWNLDIH-----IEKKRTIIPTL 272
 DB 389 KMAIDIEIATYRKLL-----EGESRISLPLPFSLSLRETNLJESLPLVDTHSKRTL---L 442
 QY 273 VEAIKEDGQREV 285
 DB 442 IKVETRDGQVIN 454

RESULT 9	ID	NAME	STANDARD:	PT:	465 AA.
1	VI	VI	VI	VI	VI
2	AC	P08667	Q51867; Q15869; Q15868;		
3	DT	01-JAN-1988	(Rel. 06, Created)		
4	DT	01-OCT-1993	(Rel. 27, Last sequence update)		
5	DT	15-JUN-2002	(Rel. 41, Last annotation update)		
6	DE	VI	VI		
7	GN	VI	VI		
8	OS	Homo sapiens (human)			
9	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
10	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
11	NCBI	TextID:9606;			
12	SE	SEQUENCE FROM N.A.			
13	MD	MEDLINE=87089701; PubMed=3467175;			
14	RA	Ferrari S., Battini R., Kaczmarek L., Rittling S., Calabretta B.,			
15	RA	de Riel J.K., Philippou V., Wei J.-F., Baasge R.;			
16	RA	"Coding sequence and growth regulation of the human vimentin gene.";			
17	MO	Mol. Cell. Biol. 6:3614-3620(1986).			
18	SE	SEQUENCE FROM N.A.			
19	MD	MEDLINE=91067467; PubMed=2251132;			
20	RA	Honore B., Madsen P., Basse B., Andersen A., Walbum E., Cells J.E.,			
21	RA	Leeffers H.;			
22	RA	"Nucleic acid sequence of cDNA covering the complete coding part of the			
23	RA	human vimentin gene.";			
24	MO	Nucleic Acids Res. 18:6692-6692(1990).			
25	SE	SEQUENCE FROM N.A.			
26	RA	Zimbelmann R.;			
27	RA	Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.			
28	SE	SEQUENCE FROM N.A.			
29	MD	Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.			
30	RA	Bird C.;			
31	RA	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
32	SE	SEQUENCE FROM N.A.			
33	MD	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
34	RA	Strausberg R.;			
35	RA	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
36	SE	SEQUENCE OF 1-134 FROM N.A., AND TISSUE SPECIFICITY.			
37	MD	MEDLINE=Breast carcinoma;			
38	RA	MEDLINE=89303836; PubMed=2472876;			
39	RA	Solmers C.L., Walker-Jones D., Heckford S.E., Worland P.,			
40	RA	Valterius E., Clark R., McCormick F., Stampfer M., Abulatach S.,			
41	RA	Gelman E.P.;			
42	RA	"Vimentin rather than keratin expression in some hormone-independent			
43	RA	breast cancer cell lines and in oncogene-transformed mammary			
44	RA	epithelial cells.";			
45	MO	Cancer Res. 49:4258-4263(1989).			
46	SE	SEQUENCE OF 112-465 FROM N.A., AND TISSUE SPECIFICITY.			
47	MD	MEDLINE=Fluoroblast;			
48	RA	MEDLINE=88226018; PubMed=3371665;			
49	RA	Perrault J., Lilienbaum A., Vasseur M., Paulin D.;			
50	RA	"Nucleic acid sequence of the human vimentin gene and regulation of its			
51	RA	transcription in tissues and cultured cells.";			
52	MO	Gene 62:7-16(1988).			
53	SE	SEQUENCE OF 166-465 FROM N.A.			
54	MD	MEDLINE=Osteosarcoma;			
55	RA	MEDLINE=90215314; PubMed=2323579;			
56	RA	Gupta A.K., Aubin J.E., Waye M.W.Y.;			
57	RA	"Isolation of a human vimentin cDNA with a long 3'-noncoding region			
58	RA	from a human osteosarcoma cell line (MG-63).";			
59	MO	Gene 86:303-304(1990).			
60	SE	SEQUENCE OF 16-24 AND 54-69.			

FT CONFLICT 277 277 S -> I (IN REF. 7).
 FT CONFLICT 338 338 S -> C (IN REF. 7).
 FT CONFLICT 349 349 N -> K (IN REF. 7).
 FT CONFLICT 441 441 F -> L (IN REF. 4, 5, 7 AND 8).
 SO SEQUENCE 465 AA; 53554 MW; 858EC8732865PFI CRC64;

Query Match 5.4%; Score 97; DB 1; Length 465;
 Best Local Similarity 20.8%; Pred. No. 0.75;
 Matches 65; Conservative 46; Mismatches 112; Indels 90; Gaps 15;

27 OEVLKRCGLRQLEPGRSRLCTLEDYEDYPPSPDPAELVLTILGAMQGYVSDIR 86
 179 EDIR---LREKLOE---MLQREAEANTLOSRQVDNLSRLRLERKVESLQBEI- 230
 87 RLFAFHPQVGLIOAA-----OQLCDQAPQORLLADLHNVSQNIATRAEDPPW 141
 231 AFLKRLHEEELQELQIQEHOVIDVDSKPDLTALRDVRYQY-ESVAKKNQEAEMW 289
 142 F-----EGLESRFQSKSGYLRY-----SCESRIR-----S 166
 290 YKSKFADLSEANNNNDALRQAKQESNEYRQVQTLCEVDALKTNESTLERQRMEN 349
 167 YLREVSYPSTVGAEQEEFLRVLSMCOPLRSQNGSYDRGAKGSRCTPEGFSC 226
 350 FAVPAANYQDTIG-RLQDEIQNMKEEMARHLREYQ-----DLIN 388
 227 QGPEPMDSCLSRHSINPYSNRESRL-----FST-----WNLDHI-----IEKRTIIP 272
 389 KMAIDIELATYRKLL-----EGESRSISLPFPFSSLNLRNRLDLSPLVDHSHKRTF---L 442
 273 VEAIKEDDGREVD 285
 443 IKVETRDGYIN 455

RESULT 10
 TIME_BOVIN STANDARD; PRT: 465 AA.

AC P48616;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vimentin.
 GN VIM.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 [1] SEQUENCE FROM N.A.
 MEDLINE-9419308; PubMed-8144034;
 HESS J.F., Casselman J.T., Fitzgerald P.G.;
 "Nucleotide sequence of the bovine vimentin-encoding cDNA."
 Gene 140:257-259(1994).
 -1- FUNCTION: Vimentins are class-III intermediate filaments found in
 various non-epithelial cells, especially mesenchymal cells.
 -1- SUBUNIT: HOMOPOLYMER.
 -1- PFM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
 MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
 DIVISION. AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
 REORGANIZED (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL, L13263; AAA53661.1; -

DR InterPro: IPR001664; IF.
 DR Pfam: PF00038; Filament. 1.
 DR PROSITE: PS00226; IF. 1.
 KW Intermediate filament; Coiled coil; Phosphorylation.
 FT INIT MET 0 0 BY SIMILARITY.
 FT DOMAIN 1 94 HEAD.
 FT DOMAIN 95 406 ROD.
 FT DOMAIN 407 465 TAIL.
 FT DOMAIN 95 130 COIL 1A.
 FT DOMAIN 131 152 LINKER 1.
 FT DOMAIN 153 244 COIL 1B.
 FT DOMAIN 245 267 LINKER 12.
 FT DOMAIN 268 406 COIL 2.
 FT MOD_RES 38
 FT MOD_RES 38
 FT MOD_RES 82
 FT MOD_RES 82
 FT PHOSPHORYLATION (BY CAMK2) (BY
 FT PHOSPHORYLATION (BY CAMK2) (BY
 FT SIMILARITY).
 FT SIMILARITY).
 SO SEQUENCE 465 AA; 53545 MW; 119E126778B5801 CRC64;

Query Match 5.4%; Score 96; DB 1; Length 465;
 Best Local Similarity 20.4%; Pred. No. 0.92;
 Matches 64; Conservative 48; Mismatches 111; Indels 90; Gaps 15;

27 OEVLKRCGLRQLEPGRSRLCTLEDYEDYPPSPDPAELVLTILGAMQGYVSDIR 86
 179 EDIR---LREKLOE---MLQREAEANTLOSRQVDNLSRLRLERKVESLQBEI- 230
 87 RLFAFHPQVGLIOAA-----OQLCDQAPQORLLADLHNVSQNIATRAEDPPW 141
 231 AFLKRLHEEELQELQIQEHOVIDVDSKPDLTALRDVRYQY-ESVAKKNQEAEMW 289
 142 F-----EGLESRFQSKSGYLRY-----SCESRIR-----S 166
 290 YKSKFADLSEANNNNDALRQAKQESNEYRQVQTLCEVDALKTNESTLERQRMEN 349
 167 YLREVSYPSTVGAEQEEFLRVLSMCOPLRSQNGSYDRGAKGSRCTPEGFSC 226
 350 FSVPAANYQDTIG-RLQDEIQNMKEEMARHLREYQ-----DLIN 388
 227 QGPEPMDSCLSRHSINPYSNRESRL-----FST-----WNLDHI-----IEKRTIIP 272
 389 KMAIDIELATYRKLL-----EGESRSISLPFPFSSLNLRNRLDLSPLVDHSHKRTF---L 442
 273 VEAIKEDDGREVD 285
 443 IKVETRDGYIN 455

RESULT 11
 DBS_HUMAN STANDARD; PRT: 1108 AA.

AC O15068;
 DT 15-JUN-1999 (Rel. 38, Created)
 DT 15-JUN-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Guanine nucleotide exchange factor DBS (DBS's big sister) (MCF2
 DE transforming sequence-like protein) (Fragment).
 GN MCF2L OR KIAA0362.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 [1] SEQUENCE FROM N.A.
 MEDLINE-9734984; PubMed-9205841;
 Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
 Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 "Prediction of the coding sequences of unidentified human genes. VII.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro."
 DNA Res. 4:141-150(1997).
 -1- FUNCTION: GUANINE NUCLEOTIDE EXCHANGE FACTOR THAT POTENTIALLY

Best Local Similarity 21.7%, Pred. No. 4.8;
Matches 93; Conservative 38; Mismatches 136; Indels 161; Gaps 21;

OY 15 SPRKFGVARSQEQE---VLRKGLR-----FOLPERG-----SLTCLYEGET 53
DB 317 APPPTAPAGAACLEKRSVASHPSCLRSITKSVLSGLPSTSOQHAFARSPACPFDKGI 376
OY 54 ---ELMEDYFSPVPD-----NAELVLTLLQAMOG----- 80
DB 377 TGGDLTKDYTPFTGNGVOPHVGOKEVSTFTGSPRLKPGLEALCKOGELEDRLRSVTFSS 436
OY 81 -----YSDIRRFISAFNHP-OGVGLIOAAOOLLCDPOAPQORLLADLL- 123
DB 437 ACQVSTSVHYSYGVSSLDKDL-EEVPRKGLMWGAGSLPSSQAYSHGIMADHLFGR 493
OY 124 ---HNVSONIAETRADDEPMEGLESRRQSGYLRSQESRIRSY-----L 168
DB 494 MRPTSCPYPIKVCRRSP-----LEIRKRTSS-----SCSS--YSTAEQSGSGSCSLPL 542
OY 169 REVSSTPSTVGA-----BAQEFRLVIGSMCQRLRSQYNGSYFDRGAKGSRILCTPEGW 223
DB 543 CEFTSSPSCQARFLATEHQBGLMGDMKNOYRPOKCEQST-----GINSDESQS 595
OY 224 FSCGPFDMSCLSRHSINPYSNBSRIEFTWNLDHIEKKRTIIPTLVEAIKPDQGRE 283
DB 596 FS---EADSESC-----PVDKQGEVKL-PEPVDQITDLPNDFOAMIKMKR----- 638
OY 284 VDWEYFYGLFTSENLKLVHTY-----CHKK-----TTKLNCDPSRI 321
DB 639 -----LTSEDLFIHDVRRSRKNRIAGORCKRKIDCIONLECRIRLVCEKREKL 688
OY 322 YKPQRLK 329
DB 689 LSERNOLK 696

RESULT 15
VIME-CHICK STANDARD: PRT: 459 AA.
AC P09654: 091023:
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vimentin.
CN VIM
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
Gallus.
NCBI_TaxID=9031;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=87250401; PubMed=3036797;
RA Zehner Z.E., Li Y., Roe B.A., Paterson B.M., Sax C.M.;
RT "The chicken vimentin gene. Nucleotide sequence, regulatory elements,
RT and comparison to the hamster gene.";
RL J. Biol. Chem. 262:8112-8120(1987).
[2]
SEQUENCE OF 453-459 FROM N.A.
RX MEDLINE=83195052; PubMed=6573660;
RA Zehner Z.E., Paterson B.M.;
RT "Characterization of the chicken vimentin gene: single copy gene
RT producing multiple mRNAs.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:911-915(1983).
CC -1- FUNCTION: Vimentins are class-III intermediate filaments found in
CC various non-epithelial cells, especially mesenchymal cells.
CC -1- STRUCTURE: HOMODIMER.
CC -1- PPM: ONE OF THE MOST PROMINENT PHOSPHOPROTEINS IN VARIOUS CELLS OF
CC MESENCHYMAL ORIGIN. PHOSPHORYLATION IS ENHANCED DURING CELL
CC DIVISION, AT WHICH TIME VIMENTIN FILAMENTS ARE SIGNIFICANTLY
CC REORGANIZED.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL: M15852; AAA49134.1; -
DR EMBL: M15850; AAA49134.1; JOINED.
DR EMBL: M15851; AAA49134.1; JOINED.
DR EMBL: V00447; CAA23726.2; -
DR PIR: A29329; A29329.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; Filament; 1.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Phosphorylation.
FT INIT_MET 0
FT DOMAIN 1
FT DOMAIN 89
FT DOMAIN 401
FT DOMAIN 89
FT DOMAIN 125
FT DOMAIN 147
FT DOMAIN 239
FT DOMAIN 262
SQ SEQUENCE 459 AA; 53011 MW; 8D1E2B588BA7B104 CRC64;
Query Match 5.1%; Score 91; DB 1; Length 459;
Best Local Similarity 19.7%; Pred. No. 2.4;
Matches 61; Conservative 49; Mismatches 115; Indels 84; Gaps 13;

OY 28 EVLRKGLRFPOLPERSRLCYEDGETELTEPFSPVDNNEVLTLTGAMGGYSDRR 87
DB 174 DIM- -LREKLEEE- -MLOREAEETLSFQDNDVNSLGLDELPVSLDEI- V 225
OY 88 FLAFLHPOVGLIOAAQ-----OLCDQAGORRLADLLHNVSONIAETRADDPWF 142
DB 226 FLKTHDEETRELQALQOHQIQIDMDYSKDLTAALRDYRQY-ESVAKKIQAEEMZY 284
OY 143 -----EGLESFQSKSYLRY-----SCESNR-----SY 167
DB 285 KSKRADLSEANNNALQAKQANEYRQIQSLTCEVDALGSMESLERQRMMEENF 344
OY 168 LREVSSTPSTVGAQAEFLRVLGSMCQRLRSQYNGSYFDRGAKGSRILCTPEGWFSQ 227
DB 345 AVEANAYQDTIG-RIODETQNKKEEMARHREYQ-----DLNVK 383
OY 228 GFPMDCLSRHSINPYSNBSRI-----LFTWNLDHIEKKRTIIP-----LVEAI 276
DB 384 MALDIEIATYRKLL-----EGESRINMPPTFASLNLRETNIESQPIVDTHSKRTLLIKTV 440
OY 277 KEQDGREVD 285
DB 441 ETRDQGVIN 449

RESULT 16
POL-STVSP STANDARD: PRT: 1022 AA.
AC P19505: 088140:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE POL polypeptide (contains: Protease (Retriopapsin) (EC 3.4.23.-);
DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)).
GN POL.
OS Simian immunodeficiency virus (PBj14/Bcl-3 isolate) (sooty mangabey).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11738;
RX NCI
[1]
SEQUENCE FROM N.A.
RX MEDLINE=90272009; PubMed=1971917;

RA Dewhurst S., Embretson J.E., Anderson D.C., Mullins J.I., Fultz P.N.;
 RT Sequence analysis and acute pathogenicity of molecularly cloned
 RT SIVSM-PB14.
 RL Nature 345:636-640(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92368737; PubMed-1503826.
 RA Dewhurst S., Embretson J.E., Fultz P.N., Mullins J.I.;
 RT Molecular clones from a non-acutely pathogenic derivative of
 RT SIVsmmPB14: characterization and comparison to acutely pathogenic
 RT clones.
 RL AIDS Res. Hum. Retroviruses 8:1179-1187(1992).
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
 CC Phosphomonester.
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate
 CC + (DNA)(N).
 CC -1- PTM: CLEAVAGE SITES THAT YIELD THE MATURE PROTEINS REMAIN TO BE
 CC DETERMINED.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
 CC KNOWN AS THE RETROPEPSIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M31325; AAA47753.1; -
 DR EMBL: 103298; AAA4777.1; -
 DR HSSP: P04584; 1PHV.
 DR HIV: M31325; POLSMMPB1.
 DR MEROPS: A02.002; -
 DR InterPro: IPR001995; Asprotease_rtrv.
 DR InterPro: IPR001969; Asprotease_sile.
 DR InterPro: IPR001037; Integrase_C.
 DR InterPro: IPR003308; Integrase_Zn.
 DR InterPro: IPR002156; Knaaseh.
 DR InterPro: IPR000477; RVTse.
 DR InterPro: IPR001584; Rve.
 DR Pfam: PF00077; naseeh. 1.
 DR Pfam: PF00077; rvt. 1.
 DR Pfam: PF00078; rvt. 1.
 DR Pfam: PF00552; Integrase; 1.
 DR Pfam: PF00665; rve; 1.
 DR Pfam: PF02022; Integrase_Zn; 1.
 DR PROSITE: PS00141; ASP_PROTEASE; 1.
 DR PROSITE: PS50175; ASP_PROT_RETROV; 1.
 KM AIDS: Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;
 Nuclease; transferase; RNA-directed DNA polymerase.
 CHAIN 1 170
 ACT_SITE 96 BY SIMILARITY.
 SEQUENCE 1022 AA; 115869 MW; 16DFEBA03F289D6A CRC64;
 Query Match 5.13; Score 90.5; DB 1; Length 1022;
 Matches Local Similarity 23.08; Pred. No. 7.6;
 Match 36; Conservative 33; Mismatches 92; Indels 63; Gaps 11;
 QY 16 PRKFGVAGRSCEVLEKRGKARFOLPERGSRCLCYEDTELETEDYFSPVNAE----LVL 71
 DB 265 PHEPAGLAKRRIRIVLDGVAFTSIP-----LDEERQRTATLTLPBV-NNAEGKRIYI 316
 QY 72 LITGQAMOG-----YSDIRFELSAFHE--POVGLIQAAQQL--CDBQAPORQL--LA 120
 DB 317 KVLPGQMKGSPALFQHMTRMNVLEPFRKANDVTLIQMDIILASDRTDLEHBRVVLQK 376
 QY 121 DLLHNTSQTAAETRAEDPPW-FEGLE-----SRFSKSGILRY 158
 DB 377 ELLNLSIGFSTPEKFKDPQFMNGYELMPTKKLQKIELDPORETVTDNIQKLYGLNW 436
 QY 159 SCE-----SRIRSYLRVSSYSTVGAEAGEEFLRVGSMQGRRLSRMQYNGSYF 207

DB 437 AAQIYPIKIKHLCRLIRGKMTLFEVQWTEMAEATYEENKI-----ILSQEGECY 489
 QY 208 DRG 210
 DB 490 QEG 492
 RESULT 17
 DBS_MOUSE STANDARD; PRT: 1149 AA.
 AC 064096;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Guanine nucleotide exchange factor DBS (DBL's big sister) (MCF2
 DE transforming sequence-like protein).
 GN MCF2L OR DBS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hematopoietic;
 RX MEDLINE-95166559; PubMed-7862449;
 RA Whitehead I., Kirk H., Kay R.;
 RT Dbs, a homolog of the Dbl guanine nucleotide exchange factor.
 RL Oncogene 10:713-721(1995).
 CC -1- FUNCTION: GUANINE NUCLEOTIDE EXCHANGE FACTOR THAT POTENTIALLY
 CC LINKS PATHWAYS THAT SIGNAL THROUGH RAC1, RHOA AND CDC42. CATALYZES
 CC GUANINE NUCLEOTIDE EXCHANGE ON RHOA AND CDC42 AND INTERACTS
 CC SPECIFICALLY WITH THE GTP-BOUND FORM OF RAC1, SUGGESTING THAT IT
 CC FUNCTIONS AS AN EFFECTOR OF RAC1. MAY ALSO PARTICIPATE IN AXONAL
 CC TRANSPORT IN THE BRAIN. BECOMES ACTIVATED AND HIGHLY TUMORIGENIC
 CC BY TRUNCATION OF THE N-TERMINUS (BY SIMILARITY)
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN SEVERAL HEMOPOIETIC
 CC CELL LINES AND IN THYMUS AND SPLEEN, AND AT HIGHER LEVELS IN OTHER
 CC TISSUES, PARTICULARLY IN BRAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 DBL-HOMOLOGY (DH) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 CRAL-TRIO DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S76838; AAB33461.2; -
 DR MGD: MGI:103263; Mcf2l.
 DR InterPro: IPR001251; CRAL_TRIO.
 DR InterPro: IPR001331; GDS_CDC24.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000219; RHOGEF.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR002017; Spectrin.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00435; spectrin; 1.
 DR Pfam: PF00621; RHOGEF; 1.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00325; RHOGEF; 1.
 DR SMART: SM00516; SEC14; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00150; SPEC; 1.
 DR PROSITE: PS50191; CRAL_TRIO; 1.
 DR PROSITE: PS00741; DH 1; 1.
 DR PROSITE: PS50010; DH 2; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.

DR PROSITE: PS50002; SH3; 1.
 KW Guanine-nucleotide releasing factor; SH3 domain; Proto-oncogene.
 FT DOMAIN 52 224 CRAL-TRIO.
 FT DOMAIN 624 627 POLY-GLU.
 FT DOMAIN 632 812 DH.
 FT DOMAIN 841 950 PH.
 FT DOMAIN 1058 1112 SH3.
 SQ SEQUENCE 1149 AA; 129148 MW; D9100DD133BEA9E3 CRC64;
 Query Match 5.1%; Score 90.5; DB 1; Length 1149;
 Best Local Similarity 22.0%; Pred No. 8.8;
 Matches 74; Conservative 41; Mismatches 134; Indels 87; Gaps 15;
 47 CLYEDTELT-EDYFSPVDNAELVLTGLGAGQGYSDIRRLSFAHEPQVGLQAAQ 105
 486 CQSDGAEAFQIEKELTGAENKIOELNIEYECILNQ-----DLLEHYQK 535
 106 LACDEAPOROLLDLH---NVSQNAIETRAEDP--PWEGL-----E 146
 536 VF-----OKQESTEMFRRQASLKLAQOTPVOPVAPREALTKSPSPSGSWSS 589
 147 SRFQSKGYLRYSCESIRSYLREY-SSYPSTVGAEAQEFVLVLSMQRISMQYNGS 205
 590 ENSSEGNALRRGPRYRAKSEMEPRQRTSGTEE--EESLAI-----LRHVMNEL 640
 206 YFDRGAKGSRLLCTPGWPGSCGPFDMQSCSRHSINPYNSRESRIFFTWMLDHIKK 265
 641 LDFERAYVEELQVLEGVAA-----EMDNPLMALISTGLQNNKNILFG--NMELIYHPH 693
 266 RTIPLTYEA--IKRQDGR-----EVDMEFYGLFTSENK 300
 694 NRRNPAGVESCIDCPPLVGRCLERMEFQYIEKCNKRPSLSMRQSDCPFOE--- 750
 301 LVHYCHKRTTKLNDPSRIYKPTRIKRPKPPK 336
 751 -----CCKLMDKRLSD-SYLLKPVORTIKYQDLKK 780
 RESULT 18
 AT06_SCHPO STANDARD; PRT; 1211 AA.
 ID AT06_SCHPO Q9UT01; Q9UT01; Q9UT01;
 AC 014072; Q9UT01; Q9UT01; Q9UT01;
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Probable cation-transporting ATPase C2E11.07C (EC 3.6.3.-).
 GN SPAC2E11.07C OR SPACUNK4.07C OR SPAPUK71.01.
 Schizosaccharomyces pombe (Pilsion Yeast).
 Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 Schizosaccharomycetales; Schizosaccharomycetaceae;
 Schizosaccharomycetes.
 NCBI_TaxID=4896;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream N.A., Lyne M., Lyne R., Stewart A.,
 RA Sgourou J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gholizadeh S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howard S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moutle S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voickeys G., Aert R., Robben J., Geymonat B.,
 RA Weltjens I., Vansteenkiste G., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer C., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Motier S.,
 RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Roche M., Galliard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,
 RA Cerutti L., Lowe T., McComble W.R., Paulsen I., Potashkin J.,
 RA Shpakowski G.V., Ussery D., Barrett B.G., Nurre P.,
 RA "The genome sequence of Schizosaccharomyces pombe";
 RT Nature 415:871-880(2002).
 RN [2]
 RC SEQUENCE OF 617-809 FROM N.A., AND SUBCELLULAR LOCATION.
 RC STRAIN=968 h50;
 RX MEDLINE=20223868; PubMed=10759889;
 RA Ding D.Q., Tomita Y., Yamamoto A., Chikashige Y., Harauchi T.,
 RA Hirooka Y.;
 RT "Large-scale screening of intracellular protein localization in living
 RT fission yeast cells by the use of a GFP-fusion genomic DNA library";
 RT genes Cells 5:169-190(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2
 CC ATPASES). SUBFAMILY V.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AL031190; CA20137.1;
 DR EMBL: AL039738; CA82214.1;
 DR EMBL: AB027853; BA867157.1;
 DR InterPro: IPR001577; ATPase_E1-E2.
 DR InterPro: IPR001454; Hlgase/Hydrlase.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 DR Pfam: PF00702; Hydrolase; 1.
 DR PRINTS: PR00119; CATALPASE.
 DR PROSITE: PS00154; ATPase_E1_E2; 1.
 KW Hydrolyase; Transmembrane; Phosphorylation; Magnesium; ATP-binding.
 DR DOMAIN 1 21
 FT TRANSMEM 22 43
 FT DOMAIN 44 49
 FT TRANSMEM 50 72
 FT DOMAIN 73 191
 FT TRANSMEM 192 214
 FT TRANSMEM 215 217
 FT TRANSMEM 218 236
 FT TRANSMEM 237 397
 FT TRANSMEM 398 417
 FT TRANSMEM 418 430
 FT TRANSMEM 431 452
 FT TRANSMEM 453 995
 FT TRANSMEM 996 1015
 FT TRANSMEM 1016 1022
 FT TRANSMEM 1023 1039
 FT TRANSMEM 1040 1057
 FT TRANSMEM 1058 1081
 FT TRANSMEM 1082 1101
 FT TRANSMEM 1102 1124
 FT TRANSMEM 1125 1135
 FT TRANSMEM 1136 1155
 FT TRANSMEM 1156 1172
 FT TRANSMEM 1173 1195
 FT TRANSMEM 1196 1211
 FT MOD.RES 485 485
 FT METAL 824 824
 FT METAL 828 828
 SQ SEQUENCE 1211 AA; 136260 MW; 6EE8E228CA5A57EC CRC64;
 Query Match 5.0%; Score 90; DB 1; Length 1211;
 Best Local Similarity 20.7%; Pred. No. 10;

Matches 64: Conservative 42: Mismatches 111: Indels 92: Gaps 12:

QY 37 FOLPERGS-RLCTYEDTELTEDYPPSPVDAEVLTLGAMOGYSDIR-----FL 89
 Db 471 FRLSLGHLIDICCDKGTGLTEEH-----MVVOGIGVNRKRDYSLEKL 514
 QY 90 SAFHEPOYGLIQAAOQLLDEQAPORORLAD-----LHNVSONIAETAEDEP---W 141
 Db 515 SDASNDALATATATHTLVLEQEGETPRKVGDPPEKATVENLGSITEKNFVSAPEGVF 574
 QY 142 FEG---LESRRPSKGYLRVSCESRIRSYLREVSSTVGAEOEFLRYLGSRCORLR 198
 Db 575 YKGRVQIIRNFQFSALKROSSVSNVRSGSFPTFSVKAP-----EVIATMLREV- 627
 QY 199 SMQYNGSYFDRGAGGSHLCTPEGTFSCQGFPMDSCLSRHSINPYNSRESRIILFTW-- 256
 Db 628 PRDEKITYKDGKRGSRVLAIGKRYFKNYIPENOVSDLSRSI-----ESDLVFAGFLI 681
 257 -----NLDH-----IIEKRTIIPITVEAIKE 278
 682 FTSPLEKEDARQTVOMLNNSSHRCAITGDNPLTAVYVAEQVGIYER-----PTVLADIKH 736
 279 QDGREVDME 287
 737 ENKILEMK 745

RESULT 19
 G6PE_RABIT
 ID G6PE_RABIT STANDARD: PRT: 763 AA.
 AC P56201;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE G6H/66GL endoplasmic bifunctional protein [includes: Glucose 1-dehydrogenase (EC 1.1.1.47) (Hexose-6-phosphate dehydrogenase); 6-phosphogluconolactonase (EC 3.1.1.31) (6PGL)].
 GN G6H.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 NC NCB1_TaxID=9986;
 RN [1]
 RP SEQUENCE.
 RC SPRAIN-New Zealand white; TISSUE=Liver;
 RX MEDLINE=93281746; PubMed=8506377;
 RA "Isolation and the complete amino acid sequence of luminal endoplasmic reticulum glucose-6-phosphate dehydrogenase.";
 RT Proc. Natl. Acad. Sci. U.S.A. 90:5302-5306(1993)
 RL
 -1- FUNCTION: OXIDIZES GLUCOSE-6-PHOSPHATE AND GLUCOSE, AS WELL AS OTHER HEXOSE-6-PHOSPHATES (BY SIMILARITY).
 -1- CATALYTIC ACTIVITY: Beta-D-glucose + NAD(P)(+) = D-glucose-1,5-lactone + NAD(P)H.
 -1- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NAD(P)(+) = D-glucose-1,5-lactone 6-phosphate + NAD(P)H.
 -1- CATALYTIC ACTIVITY: 6-phospho-D-glucose-1,5-lactone + H(2)O = 6-phospho-D-glucuronate.
 -1- COFACTOR: CAN USE BOTH NAD OR NADP.
 -1- SUBCELLULAR LOCATION: MICROSOLEM, ENDOPLASMIC RETICULUM LUMEN.
 -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE GLUCOSE-6-PHOSPHATE DEHYDROGENASE FAMILY.
 -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE ISOMERASE FAMILY.
 -1- CAUTION: THE SEQUENCE IN REF.1 SEEMS INCORRECT DUE TO THE TRANSPOSITION OF A SECTION THAT WAS PUT IN POSITION 330 TO 436 AND WHICH IS NOW TRANSPOSED TO POSITION 593 TO 700.
 DR InterPro: IPR001282; G6PD.
 DR InterPro: IPR000457; Glucosamine Iso.
 DR Pfam: PF00479; G6PD; 1.
 DR Pfam: PF01182; Glucosamine Iso; 1.
 DR Pfam: PF02781; G6PD_C; 1.

DR PRINTS: PR00079; G6PDHGRNASE.
 DR PRODOM: PD001129; G6PD; 1.
 DR TIGRFAMS: TIGR01198; pgl; 1.
 DR PROSITE: PS00069; G6P_DEHYDROGENASE; 1.
 KW Oxidoreductase; NAD; NADP; Glucose metabolism; Endoplasmic reticulum;
 KW Glycoprotein; Hydrolase; Multifunctional enzyme.
 FT MOD_RES 1 1
 FT DOMAIN 1 507
 FT DOMAIN 508 521
 FT DOMAIN 522 763
 FT ACT_SITE 189 189
 FT CARBOHYD 138 138
 FT CARBOHYD 263 263
 FT VARIANT 69 69
 FT VARIANT 75 75
 SO SEQUENCE 763 AA; 85285 MW; 24AE10390628672 CRC64;

Query Match 5.0%; Score 89.5; DB 1; Length 763;

Best Local Similarity 23.7%; Pred. No. 6.3; Mismatches 110; Indels 85; Gaps 18;

Matches 75: Conservative 47: Mismatches 110: Indels 85: Gaps 18:

QY 76 QAMOGYSDIRRLPSAFHE-POVGLIQAAOQLLDE-----QAP-QORRL 119
 Db 1 QELGHSVYI--LGATGDLAKTYLMQGLFQLFDEAGKSHSFPHGALTAAPRQGGELM 58
 QY 120 ADLHNNV--ONIAETAEDEPPWFEGLSEFROSKGYLRVSCESRIRSYLREVSSTPST 177
 Db 59 AKALESISCPRDMPALSCAE-----LQAFRLRSRYRLKTAEDYQALGRDIE----- 106
 QY 178 VGAEOEFLRVLSM-----CORLSMONG-----SYFDRGAKNG 214
 Db 107 --AOYQDGLREAGMFFTSPPRAYDIARNINSSCPGCGAMLRVYLEPFGHDLISA 164
 QY 215 SRLCTPEG-WFSCGPFDMDSCLSRHS--INPSNRESRIILFTSWNDHTIEKRTIIP 270
 Db 165 QQLATEIGSFQOEBMYRVHYLQKVAQGLTPRDQRRALDSLMNHH-VEVEVEIMK 223
 QY 271 TLVEAIKQDREVDMEFFGLL-----FISENLKY-----HYVHKRT-THKLCD 317
 Db 224 ETVVA---EGRTSFYER-YGVINDTQNLHTEITLVAMELPANVSCSEAVLRHKL--- 275
 QY 318 PSRIYKPTRLRKROPV 334
 Db 276 --QAFRALRRLRQGSAY 290

RESULT 20
 DESM_XENLA
 ID DESM_XENLA STANDARD: PRT: 458 AA.
 AC P23239;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Desmulin.
 GN DES.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus.
 NC NCB1_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90032404; PubMed=2806128;
 RA Herrmann H., Fouquet B., Franke W.W.;
 RA "Expression of intermediate filament proteins during development of Xenopus laevis. II. Identification and molecular characterization of desmin.";
 RT Development 105:299-307(1989).
 RL
 -1- FUNCTION: DESMIN ARE CLASS-III INTERMEDIATE FILAMENTS FOUND IN MUSCLE CELLS. IN ADULT STRIATED MUSCLE THEY FORM A FIBROUS NETWORK CONNECTING MYOFIBRILS TO EACH OTHER AND TO THE PLASMA MEMBRANE FROM THE PERIPHERY OF THE Z-LINE STRUCTURES.
 -1- SUBUNIT: HOMOPOLYMER.

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd

OM protein - protein search, using sw model

Run on: May 26, 2003, 12:08:32 ; Search time 55 Seconds
(without alignments)
818.886 Million cell updates/sec

Title: US-09-748-451-2
 Document: 1700

Sequence: 1 MLQPKSVKLRSPKFG.....SRIYKPTRLKRRQPVRRKQ 338

ing table: BLOSUM62

searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters:

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 08

Listing first 45 summaries

Database :

1	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1980.DAT *
2	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1981.DAT *
3	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1982.DAT *
4	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1983.DAT *
5	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1984.DAT *
6	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1985.DAT *
7	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1986.DAT *
8	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1987.DAT *
9	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1988.DAT *
10	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1989.DAT *
11	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1990.DAT *
12	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1991.DAT *
13	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1992.DAT *
14	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1993.DAT *
15	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1994.DAT *
16	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1995.DAT *
17	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1996.DAT *
18	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1997.DAT *
19	/SID5/gcgatga/genseeq/genseeqp_emb1/AA1998.DAT *
20	/SID5/gcgatga/genseeq/genseeqp_emb1/AA2000.DAT *
21	/SID5/gcgatga/genseeq/genseeqp_emb1/AA2001.DAT *
22	/SID5/gcgatga/genseeq/genseeqp_emb1/AA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1789	100.0	338	21	AAV57440	Human DNA fragment
2	1789	100.0	338	22	AAAG63552	Amino acid sequenc
3	1789	100.0	338	22	AAV72467	Human DNA fragment
4	1786	99.8	338	22	AAV41433	Caspase activated
5	1416	79.2	344	20	AAV29931	Mouse caspase-acti
6	1416	79.2	344	20	AAV29936	Mouse caspase acti
7	372	20.8	96	22	AAO09550	Human polypeptide
8	321	17.9	450	22	AAAB64406	Drosophila melanog
9	99	5.5	466	21	AAV92335	Human vimentin. H
10	99	5.5	466	22	AAAB6349	Murine vimentin.

1	1.1	99	5	466	23	ABBS57310
2	1.1	97	5	465	19	ABBS54351
3	1.1	97	5	466	21	AAB296353
4	1.3	97	5	466	22	AAB266348
5	1.4	97	5	466	23	AAU879837
6	1.5	97	5	466	23	AAU798373
7	1.6	94.5	5	1152	22	AAU798373
8	1.6	93.5	5	1120	22	AAU78855
9	1.7	93.5	5	181	22	AAU40801
10	1.8	93	5	221	22	AAU03204
11	1.9	93	5	221	22	ABG04580
12	2.0	92.5	5	214	22	AAU94441
13	2.1	92.5	5	238	21	AAU575607
14	2.1	92.5	5	238	21	AAU575607
15	2.2	92.5	5	238	21	AAU0363625
16	2.3	92.5	5	238	22	ABBS68325
17	2.4	92.5	5	266	22	AAU348868
18	2.5	91.5	5	190	20	AAU348868
19	2.6	91	5	219	21	ABBS70830
20	2.7	91	5	1719	21	ABBS70830
21	2.8	89.5	5	532	22	ABBS94422
22	2.9	89.5	5	1125	20	ABBS94422
23	3.0	85	5	142	22	AAU415527
24	3.1	85	5	150	22	AAU57444
25	3.2	85	5	31	21	AAU459429
26	3.3	85	5	31	21	AAU655939
27	3.4	85	5	31	22	AAU744758
28	3.5	88.5	4	429	22	ABP477568
29	3.6	88.5	4	465	23	ABBS77398
30	3.7	88.5	4	380	10	ABP932894
31	3.8	88.5	4	3210	9	ABP81771
32	3.9	88	4	613	22	ABP81771
33	4.0	88	4	756	21	ABP597273
34	4.1	87.5	4	666	22	ABBS67103
35	4.2	87.5	4	1035	11	AAU040283
36	4.3	87	4	331	19	AAU050823
37	4.4	87	4	311	21	AAU041212
38	4.5	87	4	366	21	AABS69191

ALIGNMENTS

RESULT 1
 AAY57440
 ID AAY57440 standard; Protein: 338 AA.
 XX
 AC AAY57440;
 XX
 DT 22-FEB-2000 (first entry)
 XX
 DE Human DNA fragmentation factor DFF40 protein sequence.
 XX
 KW Human: DNA fragmentation factor; DFF40; DFF45; apoptosis; gene therapy;
 KW cytosolic; growth; tumor.
 XX
 OS Homo sapiens.
 XX
 PN MO9954482-A1.
 XX
 PD 28-OCT-1999.
 XX
 PF 16-APR-1998: 98WO-US07895.
 XX
 PR 16-APR-1998: 98WO-US07895.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Wang X, Liu X;
 XX
 DR WPI: 2000-052702/04.
 DR N-PSDB; AA28992.
 XX
 PT DNA fragmentation factor DFF40 involved in apoptosis and related
 PT polynucleotide -
 XX

PS Claim 2; Page 127-128; 154pp; English.

XX The present sequence represents a human DNA fragmentation factor,
CC designated DFF40. Also described are: (1) a method of inducing apoptosis
CC in a cell comprising providing the cell with DFF40 which results in
CC apoptosis; (2) a method for inhibiting the growth of a cancer cell
CC comprising contacting a cancer cell with a DNA fragmentation factor
CC designated DFF40 under conditions permitting the uptake of the DNA
CC fragmentation factor by the cell where the presence of the DFF40 into
CC the cell induces apoptosis; (3) a method for treating cancer comprising:
CC (a) encoding a DFF40 DNA fragmentation factor; and (b) a promoter active
CC in the tumour cell, where the promoter is operably linked to the region
CC encoding the DNA fragmentation factor, under conditions permitting the
CC uptake of the nucleic acid by the tumour cell; (4) a method of
CC identifying a modulator of DFF40; and (5) a method of producing a
CC functional DNA fragmentation factor. An expression construct encoding a
CC DNA fragmentation factor DFF40 and DFF45 complex is provided to a cell to
CC induce apoptosis, especially in tumour cells. DFF40 is used to inhibit
CC the growth of a cancer cell, especially in humans.

XX Sequence 338 AA;

Query Match 100.0%; Score 1789; DB 21; Length 338;
Best Local Similarity 100.0%; Pred. No. 8.8e-167;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLOKPKSVKRLALSPKPKFGVAGRSCEVLRKGCRLRQLPBERGSRCLLYEDGTELTEDYF 60
DB 1 MLOKPKSVKRLALSPKPKFGVAGRSCEVLRKGCRLRQLPBERGSRCLLYEDGTELTEDYF 60
QY 61 PSYVDNAELVLTGAMOGVSDIRRFLSAFHEPOVGLIOAAQOLLCDEAPORRLA 120
DB 61 PSYVDNAELVLTGAMOGVSDIRRFLSAFHEPOVGLIOAAQOLLCDEAPORRLA 120
QY 121 DLHNVSQNTAETRAEDPWFEGLESRFQSGSYLRSCESRIRSYLREVSYPSTVGA 180
DB 121 DLHNVSQNTAETRAEDPWFEGLESRFQSGSYLRSCESRIRSYLREVSYPSTVGA 180
QY 181 EAOEFLRVLGSMQRLSMQYNSYFDRGAKGSSRLCTPEGWSCGPFDMDSCLSRHS 240
DB 181 EAOEFLRVLGSMQRLSMQYNSYFDRGAKGSSRLCTPEGWSCGPFDMDSCLSRHS 240
QY 241 INPYSNESRILFSTWMDHIEKKRTIIPLYAIEKDGREVDMEYFGLLFTSENK 300
DB 241 INPYSNESRILFSTWMDHIEKKRTIIPLYAIEKDGREVDMEYFGLLFTSENK 300
QY 301 LVHIVCHKKTTHKLNCDPSRIYKQTRLRKQPVRRKQ 338
DB 301 LVHIVCHKKTTHKLNCDPSRIYKQTRLRKQPVRRKQ 338

RESULT 2
592
MAG63592 standard; Protein: 338 AA.

AAAG63592;

DT 15-OCT-2001 (first entry)

DE Amino acid sequence of human DNA fragmentation factor 40 (DFF40).

XX Human: DNA fragmentation factor; DFF40; DFF45; apoptosis; DNase;

KW molecular chaperone; cancer cell.

XX Homo sapiens.

XX OS US2001011078-A1.

XX PN 02-AUG-2001.

XX PF 22-DEC-2000; 2000US-0748451.

XX PD 16-APR-1998; 98US-0065742.

XX (TEXA) UNIV TEXAS SYSTEM.

XX Wang X, Liu X;

XX WPI: 2001-496169/54.

DR N-PSDB; AAH74642.

XX New DNA fragmentation factor polypeptides and polynucleotides, useful
PT for inhibiting the growth of cancer cells, as well as for inducing
PT apoptosis of cells -

PS Claim 2; Page 45-46; 56pp; English.

XX The present sequence represents a human DNA fragmentation factor subunit
CC of 40 kDa, designated DFF40. The specification also describes DFF45.
CC DFF40 is capable of inducing apoptosis, and may contain a nuclear
CC localisation fragment. DFF45 acts as a molecular chaperone to direct
CC the folding of DFF40. Although all DNase activity is associated with
CC DFF40, DFF activity only occurs once DFF40 is complexed with DFF45.
CC The DFF polypeptides and polynucleotides are useful for inhibiting
CC the growth of cancer cells, and for inducing apoptosis of cells.

XX Sequence 338 AA;

Query Match 100.0%; Score 1789; DB 22; Length 338;
Best Local Similarity 100.0%; Pred. No. 8.8e-167;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLOKPKSVKRLALSPKPKFGVAGRSCEVLRKGCRLRQLPBERGSRCLLYEDGTELTEDYF 60
DB 1 MLOKPKSVKRLALSPKPKFGVAGRSCEVLRKGCRLRQLPBERGSRCLLYEDGTELTEDYF 60
QY 61 PSYVDNAELVLTGAMOGVSDIRRFLSAFHEPOVGLIOAAQOLLCDEAPORRLA 120
DB 61 PSYVDNAELVLTGAMOGVSDIRRFLSAFHEPOVGLIOAAQOLLCDEAPORRLA 120
QY 121 DLHNVSQNTAETRAEDPWFEGLESRFQSGSYLRSCESRIRSYLREVSYPSTVGA 180
DB 121 DLHNVSQNTAETRAEDPWFEGLESRFQSGSYLRSCESRIRSYLREVSYPSTVGA 180
QY 181 EAOEFLRVLGSMQRLSMQYNSYFDRGAKGSSRLCTPEGWSCGPFDMDSCLSRHS 240
DB 181 EAOEFLRVLGSMQRLSMQYNSYFDRGAKGSSRLCTPEGWSCGPFDMDSCLSRHS 240
QY 241 INPYSNESRILFSTWMDHIEKKRTIIPLYAIEKDGREVDMEYFGLLFTSENK 300
DB 241 INPYSNESRILFSTWMDHIEKKRTIIPLYAIEKDGREVDMEYFGLLFTSENK 300
QY 301 LVHIVCHKKTTHKLNCDPSRIYKQTRLRKQPVRRKQ 338
DB 301 LVHIVCHKKTTHKLNCDPSRIYKQTRLRKQPVRRKQ 338

RESULT 3

AAI72467
ID AAY72467 standard; Protein: 338 AA.

AC AAY72467;

DT 24-APR-2001 (first entry)

DE Human DNA fragmentation factor 40 (DFF40).

XX Human: DNA fragmentation factor; DFF; apoptosis; molecular chaperone;

KW gene therapy; hyperproliferative disorder; therapy; tumour; testenositis;

XX psoriasis; angiogenesis; cancer; cytosolic; neoplasia.

XX OS Homo sapiens.

XX OS US6165737-A.

XX PD 26-DEC-2000.

XX 16-APR-1998; 98US-0061702.
 XX 16-APR-1998; 98US-0061702.
 PR 16-APR-1998; 98US-0061702.

PA (TEXA) UNIV TEXAS SYSTEM.

PI Wang X, Liu X;

DR WPI: 2001-090481/10.

DR N-PSDB; AAD02498.

XX Identifying modulator of human DNA fragmentation factor 40, for
 PT treating cancer, involves contacting cell or cell-free composition
 PR comprising DFF40 with candidate substance and comparing apoptosis with
 control

Example 1; Fig 1; 52pp; English.

XX The present sequence is a human DNA fragmentation factor 40 (DFF40)
 CC capable of inducing apoptosis. DFF is a heterodimeric protein
 CC comprising 40kDa and 45kDa subunits. DFF45 acts as a molecular chaperone
 CC to facilitate the appropriate folding of DFF40 and acts as an inhibitor
 CC for DFF40. DFF40 and DFF45 are used in gene therapy. The modulators of
 CC human DFF40 activity are useful for inducing apoptosis, and for treating
 CC hyperproliferative disorders such as restenosis, psoriasis, metastatic
 CC tumours, angiogenesis and benign and malignant neoplasias. They are also
 CC used for treating cancers of the brain (glioblastoma, astrocytoma,
 CC oligodendroglioma and ependymoma), lung (liver, spleen, kidney, lymph
 CC node, pancreas, small intestine, blood cells, colon, stomach, breast,
 CC endometrium, prostate, testicle, ovary, skin, head and neck, oesophagus,
 CC bone marrow, blood, other tissue and multi-drug resistant cancer.

XX Sequence 338 AA;

Query Match 100.0%; Score 1789; DB 22; Length 338;
 Best Local Similarity 100.0%; Pred. No. 8; 8e-167;
 Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLOKPKSVKILRLSPKPKFGVAGRSQOEVLKRGCLRQLPDPERGSRCLYEDGTETLEDYF 60
 DB 1 MLOKPKSVKILRLSPKPKFGVAGRSQOEVLKRGCLRQLPDPERGSRCLYEDGTETLEDYF 60
 QY 61 PSYVDNAEVLITLGOAMOGYSDIRRFISAFHEPQVGLIOAQQILCDEQAAPORQRLA 120
 DB 61 PSYVDNAEVLITLGOAMOGYSDIRRFISAFHEPQVGLIOAQQILCDEQAAPORQRLA 120
 QY 121 DLHANSQNTAAETRAEDPPEWEGLESFRQSGYLKRSCESTRISYLRVSSYSTYGA 180
 DB 121 DLHANSQNTAAETRAEDPPEWEGLESFRQSGYLKRSCESTRISYLRVSSYSTYGA 180
 QY 181 EAOEFLRVLSGSMQRLRSMQYNGSYFDRGAKGSRCLCTPGWFSQCGPDMDSCLSRHS 240
 DB 181 EAOEFLRVLSGSMQRLRSMQYNGSYFDRGAKGSRCLCTPGWFSQCGPDMDSCLSRHS 240
 QY 241 INPYSNRESRIILFSTWMLDHIIEKKRTIIPLVLEAIKQDREVDWEYFYGLFTSENK 300
 DB 241 INPYSNRESRIILFSTWMLDHIIEKKRTIIPLVLEAIKQDREVDWEYFYGLFTSENK 300
 QY 301 LVHIVCHKTKTHKLNCDPSRIYKQTRLRKQPVARRQ 338
 DB 301 LVHIVCHKTKTHKLNCDPSRIYKQTRLRKQPVARRQ 338

RESULT 4

AA04123 standard; Protein; 338 AA.

AC AA04123;

DT 11-JUN-1999 (first entry)

DE Caspase activated nuclease CPAN.

XX Caspase activated nuclease; CPAN; apoptosis; human; heart attack;
 KW stroke; tumour.
 XX Homo sapiens.
 OS MO9910501-A1.
 XX 04-MAR-1999.
 PD 19-AUG-1998; 98MO-US17214.
 XX 06-JAN-1998; 98US-0072192.
 PR 22-AUG-1997; 97US-0056904.
 XX 27-AUG-1997; 97US-0056907

PA (CHIR) CHIRON CORP.

PI Halenbeck R, MacDonald H;

DR WPI: 1999-190620/16.

DR N-PSDB; AAX19907.

XX New isolated caspase activated nuclease - used to develop products
 PT which can inhibit or activate apoptosis for use in treating e.g.
 PT heart attack, stroke or tumours
 XX Claim 1; Page 38; 61pp; English.

XX The present sequence is a caspase activated nuclease (CPAN). Active
 CC CPAN proteins can be used for identifying compounds which inhibit
 CC apoptosis. Inhibitors of apoptosis can be used therapeutically, e.g.
 CC where cells are damaged and the apoptotic pathway is initiated such as
 CC by heart attack or stroke. The inactive CPAN proteins can be used for
 CC identifying compounds which activate apoptosis. Compounds which activate
 CC apoptosis can be used to treat tumours. The association of CPAN with a
 CC caspase-sensitive inhibitor suggests that when cells enter into
 CC apoptosis, the activation of caspases will trigger CPAN to become
 CC active.

XX Sequence 338 AA;

Query Match 99.8%; Score 1786; DB 20; Length 338;
 Best Local Similarity 99.7%; Pred. No. 1; 7e-166;
 Matches 337; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLOKPKSVKILRLSPKPKFGVAGRSQOEVLKRGCLRQLPDPERGSRCLYEDGTETLEDYF 60
 DB 1 MLOKPKSVKILRLSPKPKFGVAGRSQOEVLKRGCLRQLPDPERGSRCLYEDGTETLEDYF 60
 QY 61 PSYVDNAEVLITLGOAMOGYSDIRRFISAFHEPQVGLIOAQQILCDEQAAPORQRLA 120
 DB 61 PSYVDNAEVLITLGOAMOGYSDIRRFISAFHEPQVGLIOAQQILCDEQAAPORQRLA 120
 QY 121 DLHANSQNTAAETRAEDPPEWEGLESFRQSGYLKRSCESTRISYLRVSSYSTYGA 180
 DB 121 DLHANSQNTAAETRAEDPPEWEGLESFRQSGYLKRSCESTRISYLRVSSYSTYGA 180
 QY 181 EAOEFLRVLSGSMQRLRSMQYNGSYFDRGAKGSRCLCTPGWFSQCGPDMDSCLSRHS 240
 DB 181 EAOEFLRVLSGSMQRLRSMQYNGSYFDRGAKGSRCLCTPGWFSQCGPDMDSCLSRHS 240
 QY 241 INPYSNRESRIILFSTWMLDHIIEKKRTIIPLVLEAIKQDREVDWEYFYGLFTSENK 300
 DB 241 INPYSNRESRIILFSTWMLDHIIEKKRTIIPLVLEAIKQDREVDWEYFYGLFTSENK 300
 QY 301 LVHIVCHKTKTHKLNCDPSRIYKQTRLRKQPVARRQ 338
 DB 301 LVHIVCHKTKTHKLNCDPSRIYKQTRLRKQPVARRQ 338

RESULT 5

AA029931

AC	AA92335;
XX	
DT	21-AUG-2000 (first entry)
XX	
DE	Human vimentin.
XX	
KW	NIK1 Interacting Protein; vimentin; protein complex; cyostatic;
KW	antiviral, neuroprotective; cardiant.
XX	
OS	Homo sapiens.
XX	
PN	W0200020448-A2.
XX	
PD	13-APR-2000.
XX	
PF	06-OCT-1999; 99WO-US23314.
XX	
PR	06-OCT-1998; 98US-0167206.
XX	
PA	(CURA-) CURAGEN CORP.

Nandabalan K, Schulz VP, Yang M;
 WPI: 2000-303742/26.
 N-PSDB; AAA09307.

PT New complex of a NLK1 protein and a NLK1 protein-interacting protein,
PT useful for treating cancer, hyperproliferative disorder,
PT neurodegenerative disorder, cardiomyopathies, viral infections and
PT metabolic disorders
XX
PS
XX Example 1; Page 145-147; 172pp; English.

CC AAY92331-37 were isolated in a modified yeast two hybrid system using
CC NIK1 protein as "bait". These are known sequences which are NIK1
CC interacting proteins. The invention concerns purified complexes of a
CC NIK1 protein and a NIK1 protein-interacting protein, where the
CC interacting protein is chosen from TrkA, protein phosphatase 1alpha,
CC 14-3-3epsiloon, alpha-tropomyosin, vimentin, p0071, Inl-1, IP-1, IP-2,
CC IP-3, IP-4, or IP-5. NIK1 (also referred to as Nek2) is a human
CC homologue of the *Aspergillus nidulans* mitotic regulator, NIMA Kinase.
CC NIK1 is a serine/threonine-specific kinase and is thought to play a key
CC role in cell-cycle events leading to the onset of mitosis. The
CC complexes, their derivatives and NIK1 or NIK1-IP protein and DNA
CC sequences, etc. are useful for treating or preventing a disease or
CC disorder involving aberrant levels of the complex or protein. Such
CC disorders include cancer, hyperproliferative disorders, CC
CC neurodegenerative disorders, cardiomyopathies, viral infections and
CC metabolic disorders.

Y Match	5.5%	Score 99;	DB 21;	Length 466;
Local Similarity	20.8%	Pred. NO. 0.9;		
hes 65; Conservative	47;	Mismatches 111;	Indels 90;	Gaps 15;

```

QY 27 QEVLRKGCFLRQLEBGRSLCLTEYEGTLETEDYPSVDMNLELYLTGQAMQGVNSIR 86
Db 180 EDIMR---LRKLODE---MLQREBAMNTLQSPROVDNMSLARILDERKVESLQEEI- 231
QY 87 RFLSAFHBPQVGLIOAA-----QQLLCDEQAPORORLLADLLHNYSQNIAAETRAEDPWP 141
Db 232 AFLKRLHEETIOELQAOIQEOHQVQIDVYSKPDLTALRDYRQV-ESVAAKNQLGEABEM 290
QY 142 F-----EGLESRFPQSKSYLKY-----SCESRIR-----S 166
Db 291 YKSFADLSEANRNNDALROAKCESTETRYKOVOSLICEVDALGNTNLEKROMEMEEN 350
QY 167 YLREVSSTPYGVAEAOEFLRVLGSMQRLRSQMYNGSYFDRGAKGSGSLRCLPEGWESC 226
Db 351 FAVEAANYQDTIG-RLQDEIQNMMEEMARHLREYQ-----DLLNV 389
QY 227 QGPPDMQSCLSRHSINPYSNRESRL-----FST-----WNLDHT-----IEKKRTIIPPL 272

```

```
Dd      : . ! : : | | | | | | : | | | : | | | : |  
        390 KMALDIEIATYRLL---EGESRSILPLPNFSNLNRETNIIDSLPYDVTHSKRTL---L 443  
  
Qy      273 VEAIKEDGDREVD 285  
        : : : : | | : :  
Db      444 IKTVETRDGOVIN 456
```

RESULT 10 .
AAB66349
ID AAB66349 standard; peptide; 466 AA

AC AAB66349;

DT 05-APR-2001 (first entry)
YY

DE Murine vimentin.

KM vimentin; caspase cleavage site; apoptosis; antibody; Basedow's disease;
KM systemic lupus erythematosus; autoimmune haemolytic anaemia; AIDS; mouse

XX
XX
Mus sp.

XX
PN EP1067142-A1.

XX 10-JAN-2001.
PD

AA 07-JUL-2000; 2000EP-0305736.
PF

AA 07-JUL-1999; 99JP-0193235.
PR

PA (RIKE) RIKEN KK.

PI Morishima N, Nakanishi K, Shibata T;
 VV

DR WPI; 2001-149349/16.
XY

PT New antibody reacting with a cleavage product of vimentin but not with
PT the intact vimentin, useful for detecting apoptosis and the quantity of
PT cleavage product of vimentin or as a reagent for immunohistochemical
PT staining -
PT x

PS Disclosure; Page 12-14; 23pp; English.
 YX

CC The present invention provides an antibody which reacts with a cleavage
CC product of vimentin but not with the intact protein. This can be used to
CC detect apoptosis, which may then be used as an indicator of the
CC progression of diseases such as systemic lupus erythematosus, autoimmune
CC haemolytic anaemia, basedow's disease and acquired immunodeficiency
CC syndrome (AIDS).

Sequence 466 AA;

Query Match	5.58;	Score 99;	DB 22;	Length 466;
Best Local Similarity	20.18;	Pred. No. 0.9;		
Matches 63;	Conservative 47;	Mismatches 113;	Indels 90;	Gaps 14

```

QY      87  RLTSAFHEFQVGLTQA-----QQLLCDBQAFQORLLADLHNHSONIAETRAEDPW 141
Db      180  EDLMR---LREKQEE---MQGREASTQSFPQDDVMSLARLDLERKVESLQET- 231
QY      27  QEVLRRGCLRFQLPBEGSRILCEYEGTELTEDYFSPVDAEVLVLTLGAMQGYSDIR 86
Db      180  EDLMR---LREKQEE---MQGREASTQSFPQDDVMSLARLDLERKVESLQET- 231
QY      167  YLSEVSSYFSTGAEOEEFLVYLSCMCRLSMQYNGSYFDRGAKGSRILCEPWFSC 226
Db      291  YKSKFADLSEANRRNDLRAQKQSNERYQVOSLTCEDVLDLKGTNESLERQRMREEN 350
QY      351  FALAEANVODDTG-RLODELTOMKKEAMRHLEIYO-----DLINV 3899
Db      291  YKSKFADLSEANRRNDLRAQKQSNERYQVOSLTCEDVLDLKGTNESLERQRMREEN 350

```


Db 231 AFLKRLHEETQELQAOIQEOHVQIDVVSFKPDLTAALRDVQGY-ESVAARNLQEAWEW 289
 142 F-----EGLESRFQSKSGLYR-----SCESRIR-----S 166
 Db 290 YKSKFADISEANNNNDALROAKQESTERYKQVOSLCEVDALKTNSLERQKREMEEN 349
 167 YLREVSSTPSTVGAEOEFLRVLSMCQRLSMQYNGSYPRGAKGGSRLCTPEGWFS 226
 350 FAVEANAYQDTIG-RLQDEIONMKEEMARHLREYO-----DLINV 388
 QY 227 QGPPDMSCLSRHSINPYNSRESRL-----FST-----WNLDTI-----IEKKRTIPTL 272
 389 KMALDIEIATYRKLL---EGESRSISLPLPNFSSLNLRNRETWLDLPLVDTHSKRTF---L 442
 QY 273 VEAIKEDGPREVD 285
 443 IKVETRDGQVYN 455
 Db

RESULT 13
 AAB29635
 AAB29635 standard; Protein: 466 AA.

AAB29635;

21-FEB-2001 (first entry)

Human pollinosis-associated gene 795-encoded protein, SEQ ID NO:26.

Human: pollinosis-associated gene 795; vimentin homologue;
 IgE; immunoglobulin E; cedar pollen allergy; T-cell; reduced expression;
 detection; diagnosis; drug screening; allergic disease.

Homo sapiens.

WO200065050-A1.

02-NOV-2000.

26-APR-2000; 2000WO-JP02734.

27-APR-1999; 99JP-0120494.

(GENO-) GENOX RES INC.

(EISA) EISAI CO LTD.

Nagasu T, Sugita Y, Kashiwabara T, Oshida T, Obayashi M, Gunji S;
 Obayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K, Takahashi E,
 Yokoi A;

WPI: 2000-687343/67.

N-PSDB: AAC64226.

Pollinosis-associated gene 795 undergoing significantly low expression
 in subjects with high cedar pollen-specific IgE levels, useful in
 diagnosis of allergic diseases and screening drug candidates

Page 64-67; Claim 13; 73pp; Japanese.

The invention relates to the human pollinosis-associated gene 795 which
 exhibits significantly reduced expression in the T-cells of individuals
 with high cedar pollen-specific IgE (immunoglobulin E) levels. The gene
 was isolated from T-cells from individuals allergic to cedar pollen using
 the differential display method. Pollinosis-associated gene 795 has
 homology with the human vimentin gene. The invention also relates also
 relates to the protein encoded by pollinosis gene 795; to expression
 constructs and host cells comprising pollinosis-associated gene 795
 nucleic acids; pollinosis-associated gene 795 primers and probes;
 antibodies against the protein encoded by the gene; methods of detection
 of pollinosis-associated gene 795 nucleic acids; and a method of
 diagnosis of allergic diseases via the detection of pollinosis-associated
 gene 795 nucleic acids. The invention additionally encompasses methods of

CC screening drug candidates for the treatment of allergic disease by
 CC measuring the expression of pollinosis-associated gene 795 in pollen
 CC antigen-stimulated T-cells in the presence of a test compound relative to
 CC a control. Pollinosis-associated gene 795 is useful in the diagnosis of
 CC allergic diseases and in the screening of drug candidates for the
 CC treatment of such diseases. The present sequence represents a
 CC protein encoded by human pollinosis-associated gene 795.

SQ Sequence 466 AA;

Query Match 5.4%; Score 97; DB 21; Length 466;

Best Local Similarity 20.8%; Pred. No. 1.4; Mismatches 112; Gaps 15;

Matches 65; Conservative 46; Indels 90;

27 QEVLRKGLRQLEPGRGSRCLYEDGETLTDYPPVSDNAELVLTGQAMQGVSDIR 86
 180 EDIMR---LRKLEOE---MLQREEMNTIQSRQVDNASTARLDERVEVLEQERT 231

87 RFLSAFHEPQVGLQAA-----QQLCDEQAPQKORLLADLHNVSQNIATRAEDPPW 141

232 AFLKRLHEETQELQAOIQEOHVQIDVVSFKPDLTAALRDVQGY-ESVAARNLQEAWEW 290

142 F-----EGLESRFQSKSGLYR-----SCESRIR-----S 166

291 YKSKFADISEANNNNDALROAKQESTERYKQVOSLCEVDALKTNSLERQKREMEEN 350

167 YLREVSSTPSTVGAEOEFLRVLSMCQRLSMQYNGSYPRGAKGGSRLCTPEGWFS 226

351 FAVEANAYQDTIG-RLQDEIONMKEEMARHLREYO-----DLINV 389

227 QGPPDMSCLSRHSINPYNSRESRL-----FST-----WNLDTI-----IEKKRTIPTL 272

390 KMALDIEIATYRKLL---EGESRSISLPLPNFSSLNLRNRETWLDLPLVDTHSKRTF---L 443

273 VEAIKEDGPREVD 285

444 IKVETRDGQVYN 456

RESULT 14

AAB6348
 ID AAB6348 standard; peptide: 466 AA.

AAB6348;

05-APR-2001 (first entry)

Human vimentin.

Vimentin; caspase cleavage site; apoptosis; antibody; Basedow's disease;
 systemic lupus erythematosus; autoimmune haemolytic anaemia; AIDS; human.

Homo sapiens.

EPI067142-A1.

10-JAN-2001.

07-JUL-2000; 2000EP-0305736.

07-JUL-1999; 99JP-0193235.

(RIKE) RIKEN KK.

Morishima N, Nakanishi K, Shibata T;

WPI: 2001-149349/16.

New antibody reacting with a cleavage product of vimentin but not with
 the intact vimentin, useful for detecting apoptosis and the quantity of
 cleavage product of vimentin or as a reagent for immunohistochemical
 staining

XX 05-FEB-2001; 2001WO-US04098.
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX (HSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK52970.
 DR Nucleic acids encoding polypeptides with cytokine-like activities,
 useful in diagnosis and gene therapy -
 PS Claim 20; Page 353-354; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM7833-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX Sequence 1152 AA;
 SO
 Query Match 5.38; Score 94.5; DB 22; Length 1152;
 Best Local Similarity 21.28; Pred. No. 9.4;
 Matches 83; Conservative 54; Mismatches 155; Indels 99; Gaps 18;
 PS 4 KPSVKRLAL-----RSPKFGVAGRSQ-----EYLKRGCLR--FQDPERGSRCLTYE 50
 DB 489 RPKQGEHRHLCDFSAEIAIRRGILSKSLHRRLETSMKWCDEGTYLLASQPYDKCQSQ 548
 CC 51 DGELELT-EDYFPPVPDPAELVLLTLGAMQGVSDIRRLT-----SAFHE 94
 CC 549 DGAEEALQETKELEFETGAENKIDELNAYKEYESILNQDIMEHVAKVFOQASMEVEFHR 608
 CC 95 POGVL--IQAAQQLCDEQAPQORRLADLLHNVSQIAETAEDPPEWEGLESRRQSK 152
 DB 609 RQSLKTLARQTRPVQVAPREPAL-----AKSPSPSGIRGSGNS--SSE 654
 CC 153 SGYLRYSCEGRIRSYLREVSSTVGAEOEFLVYIGS--MQRIRSMQYNSYDRG 210
 DB 655 GGAIRRPYRRAKSEMSESRQGRSAGE--EESLAILRRHWSELDTER--AYVE-- 707
 CC 211 AKGGSRLCTPEGFGSCGPPDMDCSLRHSINPYSNRESRLFTSTWNLDIHIE----- 263
 DB 708 -----ELLCVLEGYAA-----EMDNPLMAHLSTGLHNKKVYLG--NMEEIYHNNRIFL 756
 CC 264 ----KKRIITPLV-----EAIKEDGREGVMEYFGLLFTSENKLVHIV 305
 DB 757 RELENTDCEPLVGRCEFLRMEDEFOIYEKCONKPRSESIMROCSDCPFOE----- 808
 CC 306 CHKKTTHKNCDSRIYKPPOTRLKROPVARK 336

DB 809 CQRKDHKLSD-SYLKPVQRTKTYOLLK 838
 RESULT 17
 AAM78853
 ID AAM78853 standard; Protein; 1120 AA.
 XX
 AC AAM78853;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1515.
 XX
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 XX WO200157190-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001WO-US04098.
 XX
 XX 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 XX (HSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR N-PSDB; AAK51986.
 DR Nucleic acids encoding polypeptides with cytokine-like activities,
 useful in diagnosis and gene therapy -
 PS Claim 20; Page 3811-3813; 6221pp; English.
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM7833-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX Sequence 1120 AA;
 SO
 Query Match 5.28; Score 93.5; DB 22; Length 1120;
 Best Local Similarity 21.28; Pred. No. 11;
 Matches 83; Conservative 54; Mismatches 155; Indels 99; Gaps 18;
 PS 4 KPSVKRLAL-----RSPKFGVAGRSQ-----EYLKRGCLR--FQDPERGSRCLTYE 50
 DB 457 RPKQGEHRHLCDFSAEIAIRRGILSKSLHRRLETSMKWCDEGTYLLASQPYDKCQSQ 516

```

OY 51 DGETLT-EDYEPSPDNALVLTLLGAMOGYVSDIRRL-----SAFHE 94
Db 517 DGAELAQEIEKFELETGEMKIOELNAYEYSIIMODLMEHVAKVFOKQASMEYFHR 576
OY 95 PGVGL--IOAAQOLICDROAPORORLADLHNVSONIAETRAEDPMPFGLSRRQSK 152
Db 577 RQASLKKLAAROTRVPQVAPRPEAL-----AKSPCSPICRGSENS--SSE 622
OY 153 SGYLRSCSRISRYLSRVSSYSTVAEAOEFLVIGS--MCORLSMOYNGSYDRG 210
Db 623 GGLLRNGPYRAKSSVSSRQSGSAGEE--EESLAILRHVMSLDTLR--AYVE-- 675
OY 211 AKGSRICPEECMEFCOCPFMDSCLSRHSINYSRRESRIILFTWNLDIIE----- 263
676 ----ELICVLEGVYA-----EMDNPPLMAHLSTGLHKKQVLEG--NMEEIYHFRNRIFL 724
264 ----KKRTITPILV-----EAIKQDQREVDMETFYGLLFTSENLKLVHIV 305
725 RELENTDCPELVGCGFLERMEDEQIYEKTCQNKPRSESLMQCSDCPFPD----- 776
OY 306 CHKKTTHKLNCDPSRIYKPOTRLKRPVK 336
Db 777 CORKIDHKLSD-STYLLKPVQRIIRYQLLKK 806

RESULT 18
AAM40801
ID AAM40801 standard; Protein: 181 AA.
AC AAM40801;
XX
XX 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 5732.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukemia.
XX
XX Homo sapiens.
XX
XX W0200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000MO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
XX
XX 25-APR-2000; 2000US-0552317.
XX
XX 09-JUL-2000; 2000US-0598042.
XX
XX 19-JUL-2000; 2000US-0620312.
XX
XX 03-AUG-2000; 2000US-0653450.
XX
XX 14-SEP-2000; 2000US-0662191.
XX
XX 19-OCT-2000; 2000US-0693036.
XX
XX 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,
XX Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX
XX N-PSDB; AA159957.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
XX
XX Example 2; SEQ ID NO 5732; 10078bp; English.

```

```

XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
XX the encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX CC localised neuropathies and central nervous system diseases, such as
XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX CC utilisation of the activities such as: Immune system suppression,
XX CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX CC assays for receptor activity, arthritis and inflammation, leukemias and
XX CC C.N.S. disorders.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification.
XX
XX Sequence 181 AA:
XX
XX
XX Query Match 5.2%; Score 93; DB 22; Length 181;
XX Best Local Similarity 24.5%; Pred. No. 0.86;
XX Matches 50; Conservative 22; Mismatches 56; Indels 76; Gaps 11;
OY 24 RSCOEVLK-----GCLRFOLPERGSRILCYEDGETLT 56
Db 4 RSCQSPRRKSRRAHVTWLVCGFTSFSLPLYICGLRF--PE--RTCSQLQADWA 58
OY 57 EDYFSP--VPDNALVLTLLGAMOGYVSDIRRLSFHPQVGIIOAAQOLICDROAPQ 114
Db 59 PDFGSPSPV-----SWGATRTGARKRLAFN--INLTGTREQ----- 94
OY 115 RQRLADLHNVSONIAETRAEDPMPFGLSRRQSGYLRSCSRISRYLR--EVS 172
Db 95 -----AHRIALNLRQSGKQDP--GLTKVYGIGYIDENKRLQAVSTNILLDFEVT 143
OY 173 S-----YPTVGAEOEFLVIGS 192
Db 144 ALHTVYETIC-REAOELSLPVVGS 166

RESULT 19
AAM03204
ID AAM03204 standard; Protein: 221 AA.
XX
XX AAM03204;
XX
XX 12-SEP-2001 (first entry)
XX
XX Human CIDE protein consensus sequence.
XX
XX HLIAL47; CIDE; Cell death inducing DFF45-like effector;
XX apoptosis; programmed cell death; cancer; haematological disorder;
XX bone marrow failure; myelodysplastic syndrome; aplastic anaemia;
XX KW neutropaenia; leukemia; cardiovascular disease; ischaemia;
XX KW reperfusion; liver disease; inflammation; neurodegenerative disease.
XX
XX Synthetic.
XX
XX OS Homo sapiens.
XX
XX Mus sp.
XX
XX
XX Key Location/Qualifiers
XX
XX FT Misc-difference 70 /label= OTHER
XX
XX FT /note= "Other- unknown"
XX
XX FT Misc-difference 124 /label= OTHER
XX
XX FT /note= "Other- unknown"
XX
XX FT Misc-difference 173 /label= OTHER
XX
XX FT /note= "Other- unknown"
XX
XX FT Misc-difference 185 /label= OTHER
XX
XX

```

FT /note- "Other- unknown"
FT Misc-difference 213
FT /label- OTHER
FT /note- "Other- unknown"
FT Misc-difference 215..217
FT /label- OTHER
FT /note- "Other- unknown"
FT Misc-difference /label- OTHER
FT /note- "Other- unknown"
PN EPI103608-A1.
XX 30-MAY-2001.
XX 25-NOV-1999; 99EP-0402926.
XX 25-NOV-1999; 99EP-0402926.
XX (SNFI) SANOFI-SYNTHELABO.
PA Grallhe P, O'Brien D;
WPI: 2001-357955/38.
XX New HIVAL47 polypeptides and polynucleotides, useful for treating
PT cancer, hematological disorders, bone marrow failure such as
PT myelodysplastic syndromes, aplastic anemia, neutropenia and leukemia,
PT cardiovascular diseases -
XX Disclosure; Fig 2; 31pp; English.
XX The sequence represents a CIDE (cell death inducing DFF45-like effector)
CC consensus sequence, thought to be involved in apoptosis and programmed
CC cell death. Human HIVAL47 polypeptides (a CIDE B protein) and
CC polynucleotides are useful in the treatment of cancer, hematological
CC disorders, e.g., bone marrow failure including myelodysplastic syndromes,
CC aplastic anemia, neutropenia and leukemia, cardiovascular
CC diseases (e.g. ischemia and reperfusion), liver diseases, inflammation
CC and neurodegenerative pathologies. These may also be used as research
CC reagents and materials for the discovery of treatments and diagnostics
CC for animal and human diseases, for chromosome identification, as
CC immunogens to produce HIVAL47 specific antibodies, and in screening for
CC (ant)agonists of the HIVAL47 polypeptide.
XX Sequence 221 AA;
SQ
Query Match 5.2%; Score 93; DB 22; Length 221;
Best Local Similarity 40.9%; Pred. No. 1.2;
Matches 27; Conservative 9; Mismatches 28; Indels 2; Gaps 2;
14 RSPRRFGVAGRSCQEVLRKGLRQLPERSRLCYEDGTEL-TEDYFPSPVDAEVLVL 72
45 RSSRK-CVTAAASIQELLSTLDLVLTGVTAVLEBDGQAVASEDFQLEDDTHLMVL 103
73 TLGOAW 78
104 EKGQSW 109
RESULT 20
ABG04580
ID ABG04580 standard; Protein; 214 AA.
XX
AC ABG04580;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #4571.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX

OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX N-PSDB; AAS68767.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
XX Claim 20; SEQ ID NO 34939; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 214 AA;
SQ
Query Match 5.2%; Score 92.5; DB 22; Length 214;
Best Local Similarity 33.3%; Pred. No. 1.2;
Matches 33; Conservative 16; Mismatches 41; Indels 9; Gaps 4;
5 PKSVKLR-----ALRSPRRFGVAGRSCQEVLRKGLRQLPERSRLCYEDGTEL-TE 57
47 PKAPRRPGRVSTADRSVRK-GIMAYSLIEDLLVDRDTLMADPFLVLEDDGTIVETE 105
58 DYFSPVPMNAELVLLTGOAWQ-GYVSDIRRFSLAFHEP 95
106 EYFOALAGDITVEMVLQKQKMPSESGTRRPLSLSKRP 144
Search completed: May 26, 2003, 15:21:03
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 15:21:43 ; Search time 3125 Seconds
(without alignments)
3147.759 Million cell updates/sec

Title: US-09-748-451-2
Perfect score: 1789
Sequence: 1 MDAQKSVKALALSPKRF.....SRIYPRRIKRPVRRKQ 338

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODER-frame+ p2n.model -DEV-xml
-O=/gen2.1/USPTO.spool/US09748451/runat_21052003.153829.16279/app_query.fasta_1.519
-DB=GenEmbl -OPM=firstap -SUPER=irge -MINMATCH=0.1 -LOOPEXT=0
-INITS=bits -START=1 -END=1 -MATRIX=biolum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=opt -THR MAX=100 -THR MIN=0 -ALIGN=20 -MODE=LOCAL
-CURRENT=1 -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09748451 -GCN 1.1 2496 -runat_21052003.153829.16279 -NCPU=6 -ICPU=3
-NO_MAP -LARGEROVER -NEG_SCORES=0 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-DEV -TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FEAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
GenEmbl :
1: gb_ba :
2: gb_hgt :
3: gb_in :
4: gb_cm :
5: gb_ov :
6: gb_pat :
7: gb_ph :
8: gb_pl :
9: gb_pr :
10: gb_ro :
11: gb_sts :
12: gb_sy :
13: gb_un :
14: gb_vi :
15: em_ba :
16: em_fun :
17: em_hum :
18: em_in :
19: em_mu :
20: em_om :
21: em_or :
22: em_ov :
23: em_pat :
24: em_ph :
25: em_pl :
26: em_ro :
27: em_sts :
28: em_un :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1789	100.0	1017	9	AB013918 Homo sapi
2	1789	100.0	2839	6	AR122560 Sequence
3	1789	100.0	2839	9	AF064019 Homo sapi
4	1786	99.8	1017	9	AF039210 Homo sapi
5	1759.5	98.4	2926	9	AB028913 Homo sapi
6	1758.5	98.3	2926	9	AB028911 Homo sapi
7	1757	98.2	1028	9	AF409060 Homo sapi
8	1754.5	98.1	1043	9	AF409061 Homo sapi
9	1743	97.4	1071	9	AF409062 Homo sapi
10	1730	96.7	3008	6	AB028912 Homo sapi
11	1416	79.2	1038	6	E36548
12	1416	79.2	1038	6	E36548
13	1416	79.2	1038	6	E36548
14	1392	77.8	1316	10	AB009377 Mus muscu
15	1251	69.9	15585	2	AL355811 Homo sapi
16	1065	58.3	1459	5	AF406761 Gallus ga
17	1043	52.6	1159	5	AF286179 Danto rer
18	941	22.9	160716	2	AL691523 Homo sapi
19	410	19.9	1833	9	AK098413 Homo sapi
20	355.5	18.4	4339	5	AF426316 Danto rer
21	321	17.9	2118	3	AF149797 Drosophi
22	320	17.9	1401	3	AB036773 Drosophi
23	320	17.9	211505	2	AL806525 Mus muscu
24	295	16.5	197159	3	AC009340 Drosophi
25	295	16.5	261432	3	AE003639 Drosophi
26	294	16.4	18765	2	AC012906 Drosophi
27	124.5	7.0	1851	9	HSV1MENTA
28	111.5	6.2	1600	9	BC030573
29	109.5	6.1	59280	10	AL603906
30	108.5	6.1	6943	8	MM005823
31	108.5	6.0	1438	10	PSA22156
32	108	6.0	1165	10	AF136601
33	106.5	6.0	1165	10	AE002394
34	106.5	6.0	11645	2	LMFCHCR36.22
35	106.5	6.0	110000	2	LMFCHCR36.22
36	106.5	6.0	172148	3	LMFCHCR36.22
37	106.5	6.0	195767	1	NMA72491
38	106.5	6.0	349980	6	AX044029 Sequence
39	105.5	5.9	1714	8	AF032688 Oryza sat
40	105.5	5.9	184765	2	AC129718 Mus muscu
41	104.5	5.8	161223	2	AC124011 Mus muscu
42	104	5.8	7117	4	PICAPOLPB7
43	104	5.8	7117	4	PICAPOLPB7
44	103	5.8	963	9	HSCIC68CC
45	102.5	5.7	1227	9	HUMVIM3

RESULT 1

ALIGNMENTS

Deleted
PRI 11-AUG-1998

AB013918
LOCUS AB013918 1017 bp mRNA linear
DEFINITION Homo sapiens mRNA for CAD, complete cds.
ACCESSION AB013918
VERSION AB013918.1 GI:3410908
KEYWORDS CAD.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (sites)
REFERENCE
AUTHORS Mukae, N., Enari, M., Sakahira, H., Fukuda, Y., Inazawa, J., Toh, H. and Nagata, S.
TITLE Molecular cloning and characterization of human caspase-activated DNase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (16), 9123-9128 (1998)
MEDLINE 98356121
REFERENCE 2 (bases 1 to 1017)
AUTHORS Nagata, S.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1998) Shigekazu Nagata, Osaka University Medical School, B-3, Department of Genetics; 2-2 Yamada-oka, Suita, Osaka 565-0871, Japan (E-mail: nagata@genetic.med.osaka-u.ac.jp, Tel: 81-6-879-3310, Fax: 81-6-879-3319)
FEATURES
SOURCE Location/Qualifiers
1..1017
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="1p36.3"
/cell_line="KT-3"
/gene="CAD"
1..1017
/gene="CAD"
1..1017
/gene="CAD"
/codon_start=1
/product="CAD"
/protein_id="BAA32250.1"
/db_xref="GI:3410908"
/translation="MLQPKSVKLRALSPKFGVAGRSCEVLRKGLRQLPERSGRLCYDEGTETLETPSPVDPNNAELVLTQAGVYSDIRRFISAHHEPQVIGTIOAOOLCDAPORORRLADLHNSONIAETRAEDPMPFEGISRSRSGSYLRISCESRTRSTREVSSTPYTGAEAGQEFELVLSKORLSMOTNGSYFRGKAGSRRLCTPEGWFSCQGPDMDSLRHSINPYSRRESRIILFTWNLDHIIKKRTIIPTLVEAIKEQGRVEMWEIFYGLFETSNLKLIVHCHKKTTHKLNCDPSRIYKPTRLKROPYRKRO"

BASE COUNT 236 a 293 c 303 g 185 t
ORIGIN

Alignment Scores:
1.9e-159 Length: 1017
1789.00 Matches: 338
100.00% Conservative: 0
100.00% Mismatches: 0
100.00% Indels: 0
Gaps: 0

US-09-748-451-2 (1-338) x AB013918 (1-1017)

QY 1 MetLeuGlnLysProGlnSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
DB 1 ATGTCTCCGAAGCCCAAGAGCGTGAAGCTGGCGGCCCGCCAGGCCAGGAAGTTCGCG 60
QY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
DB 61 GTGGCTGGCCGAGAGTGGCAGAGTGGTGGCGCAAGGCTGTCTCCGCTTCCAGCTCCCT 120
QY 41 GlnArgGlySerArgLysLeuArgLysGlnLysPheGlnLeuArgLysPheGlnLysPhe 60
DB 121 GAGCGCGGCTCCCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 180
QY 61 ProSerValProAspAlaGlnLeuValLeuLeuThrLeuGlnAlaTrpGlnGly 80

DB 161 CCGACGTGTCCCGAGACAGCCGAGCTGTGTCTGCTACCTGGGCGAGCCGTGGAGGCGC 240
QY 81 TyrValSerAspAlaArgArpPheLeuSerAlaPheHisGluProGlnValGlyLeuLeu 100
DB 241 TATGTAGAGACATGAGGCGCTTCTCCTCAGTGATTTTCAGAGCCACAGATGGGCTCATC 300
QY 101 GlnAlaAlaGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
DB 301 CAGCGCGCCAGACACCTGTGTGTATGAGAGGCGCCACAGAGCAGAGCGTGTGTGCT 360
QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspProPro 140
DB 361 GACCCCTCGACACAGCTGACGAGCAATCGGGCGGAGCCGCGGCTGAGAGACCCGCG 420
QY 141 TrpPheGlnGlyLeuGlnSerArgPheGlnSerIleValSerGlyTyrLeuArgTyrSerCys 160
DB 421 TGGTTGAAGCTTGGAGTCCCGATTTACAGACCAAGCTTGGCTATCTAGATACGCTGT 480
QY 161 GluSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
DB 481 GAGACCGGAGACCGGAGTTACCTAGAGGAGGTGAGCTCTACCCCTCCAGCGTGGCTCG 540
QY 181 GlnAlaGlnGlnGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
DB 541 GAGGCTCAGAGGAATTCCTGGGCTCTCGGCTCTCAGTGCAGAGGCTCGGCTCATG 600
QY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaLysGlyLysArgLeuCysThrPro 220
DB 601 CAGTCAATGGCAGCTTACTTGCACAGAGAGCAAGGCGGAGCGGCTCTGACACCG 660
QY 221 GlnGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
DB 661 GAAGCGTGTCTCTCTGCGAGGCTCCCTTGTGACATGAGCAGAGTGTATCAAGACATCC 720
QY 241 IleAsnProTyrSerAsnArgGluSerArgIleLeuPheSerThrTyrAsnLeuAspHis 260
DB 721 ATCAACCCCTCAGTACAGAGAGGAGAGGAGGATCTCTTACAGACCTGAGATCAG 780
QY 261 IleIleGlnLysLysArgThrIleLeuProThrLeuValGluAlaIleLysGlnGlnAsp 280
DB 781 ATAAATAGAAAAGAAAGCAACCATCTATCTCTACACTGTTGGAAGCAATTAAGAACAAAT 840
QY 281 GlnArgGluValAspTrpGlnLysArgPheTyrGlyLeuLeuPheThrSerGlnAsnLeuLys 300
DB 841 GGAAGAGAGAGTGGAGTGGAGATTTATGCGCTGCTTTTACCTCAGAGAACTATAAA 900
QY 301 LeuValHisIleValCysHisLysArgThrThrHisLysLeuAsnCysAspProSerArg 320
DB 901 CTAGTGCACATTTGCTGCCATTAAGAAACCAACCAAGCTCACTGATGACCCAGACAGA 960
QY 321 IleTyrLysProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
DB 961 ATCTCAAAACCCAGACAAAGTTGAAGCGGAAGCAGCTGTGGGAAACGCCAG 1014

RESULT 2
AR122560 2839 bp DNA linear PAT 16-MAY-2001
LOCUS AR122560
DEFINITION Sequence 1 from patent US 6165737.
ACCESSION AR122560
VERSION AR122560.1 GI:14106877
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2839)
AUTHORS Wang, X. and Liu, X.
TITLE DNA fragmentation factor involved in apoptosis
JOURNAL Patent: US 6165737-A 1 26-DEC-2000;
FEATURES 1..2839 Location/Qualifiers
source /organism="unknown"
BASE COUNT 644 a 701 c 753 g 739 t

ORIGIN

Alignment Scores:

Pred. No.: 7,09e-159 Length: 2839
 Score: 1789.00 Matches: 338
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-748-451-2 (1-338) x AR122560 (1-2839)

QY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
 132 ATGCTCCAGAACCCCAAGACGTCGAGGCGCCGCGCCACCGGAGAGTTCGCGC 191
 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 192 GTGGTGGCCGAGACTCCAGAGAGTGGTGGCCAGAGGCTGTCTCCGCTCCAGCTCCCT 251
 QY 41 GluArgGlySerArgLysCysLeuArgLysGlnValPheGlyThrGluLeuArgLysPhe 60
 252 GAGCGCGGTTCGCGGCTGGCTGCTGACGAGTGGACGAGCTGACGAGAGATTAATTC 311
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaArgPheGly 80
 312 CCCAGTGTCCCGCAACAGCGCGAGTGTGTGCTGCTCACCCTTGGCGCGAGGCTGCAAGGCG 371
 QY 81 TyrValSerAspLysArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuLeu 100
 372 TATGTGAGCGACATCAGCGCTTCTCTGATGATTCACGAGCCACAGGTGGGCTCATC 431
 QY 101 GlnAlaGlnGlnGlnLeuLeuCysAspGlnGlnAlaProGlnArgGlnArgLeuAla 120
 432 CAGGCGCGCGCGAGCTGTGTGTGATGAGCAGCGCCACAGAGCGAGGCTGTGCTGCTGCT 491
 QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaGluThrArgAlaGluAspProPro 140
 492 GACCTCTGCAACAGTCACGCAAGAACATCGGCGCGAGACCGCGGTGAGAGCCCGCGC 551
 QY 141 TrpPheGlnGlyLeuGlnSerArgPheGlnSerLysSerGlyTyrLeuArgTyrSerCys 160
 552 TGGTTGAAGCTTGGAGTCCCATTTTCAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
 QY 161 GluSerArgLysLeuArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
 612 GAGAGCGCGAGTCCGAGTCTGAGGAGGAGTACGCTGACCCCTCCACAGTGGTGGC 671
 QY 181 GluAlaGlnGlnGlnLeuLeuValLeuGlySerMetCysGlnAlaGluArgSerMet 200
 672 GAGGCTCAGAGGAATTCCTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
 QY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaLysGlyGlySerArgLeuCysThrPro 220
 732 CAGTCAATGCGACACTCTCGACAGAGGAGCCAGAGGCGGAGCGGCTCTGCAACCG 791
 QY 221 GluGlyTrpPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 792 GAAGCGGTGCTCTCTGCGAGGCTCTTGCATGAGACAGCTGTATCAAGACACTCC 851
 QY 241 IleAsnProTyrSerAspArgGluSerArgIleLeuPheSerThrTyrAsnLeuAspHis 260
 852 ATCAACCCCTACAGTAACAGAGAGAGAGATCTCTTACAGACCTGGAACTGGATGATCAC 911
 QY 261 IleIleGlnLysArgThrIleIleProThrLeuValGlnAlaIleLysGlnGlnAsp 280
 912 ATATATGAAAGAAAGAACGACATATTCCTACACCTGTGGAGAAATTAAGAACAAAGAT 971
 QY 281 GlyArgGluValAspTrpGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLys 300
 972 GGAAGAGAGTGGAGTATTTTAAAGGCTGCTTAACTCCACAGAACCTATAAA 1031
 QY 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnProSerArg 320

Db 1032 CTATGTCACATTTCTCTCCATTAAGAAACACCCACAGCTCACTGTGACCGAGACAGA 1091
 QY 321 IleTyrLysProGlnThrArgLeuArgLysGlnProValArgLysArgGln 338
 Db 1092 ATCTACAAACCCCGACAGACGTTGAGAGCGAGAGCGCTGTGCGGAAAGCGCAG 1145
 RESULT 3
 AF064019
 LOCUS
 DEFINITION
 Homo sapiens DNA fragmentation factor 40 kDa subunit (Pff40) mRNA,
 complete cds.
 ACCESSION
 AF064019
 VERSION
 AF064019.1
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2839)
 The 40-kDa subunit of DNA fragmentation factor induces DNA
 fragmentation and chromatin condensation during apoptosis
 Proc. Natl. Acad. Sci. U.S.A. 95 (15), 8461-8466 (1998)
 86357957
 MEDLINE
 PUBMED
 9671700
 REFERENCE
 2 (bases 1 to 2839)
 Liu, X., and Wang, X.
 Direct Submission
 Submitted (07-MAY-1998) Biochemistry, University of Texas
 Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd.,
 Dallas, TX 75235, USA
 FEATURES
 source
 1..2839
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..2839
 /gene="Pff40"
 132..1148
 /function="nuclease; induces DNA fragmentation and
 chromatin condensation during apoptosis"
 /codon_start=1
 /product="DNA fragmentation factor 40 kDa subunit"
 /protein_id="AAC39920.1"
 /db_xref="GI:3347857"
 /translation="MLOKPSVYLRALSPRRGVAGSGCOEVLKGLRPOLESGS
 RCIYEDPTETEDYPSVDNAEYLITLGOAGGYSDIRRELSAPRPQGLTGA
 AQLICDQARORRLADLIHNSONTIAFRADPPVDSIETRESROSGYLRSC
 ESRINSLREYSIPSTYVGENDEERLUSGRCRISNOYNSYDRCAKGSRLC
 IPEKFSQGFEDMSLSRHSINPISNESRLISTWINDITTEKRTIIPYVPAI
 KQDSEKVDMEYFYGLFTSENLKIVHIVCKRTTKLNCDSKRTIYQTRLRKQPV
 RKRO"
 BASE COUNT 644 a 701 c 755 g 739 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,09e-159 Length: 2839
 Score: 1789.00 Matches: 338
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-748-451-2 (1-338) x AF064019 (1-2839)
 QY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
 132 ATGCTCCAGAACCCCAAGACGTCGAGGCGCCGCGCCACCGGAGAGTTCGCGC 191
 QY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 192 GTGGTGGCCGAGACTCCAGAGAGTGGTGGCCAGAGGCTGTCTCCGCTCCAGCTCCCT 251

QY 41 GluArgLysSerArgLeuCysLeuTyrGluAspGlyThrGluLeuThrGluAspTyrPhe 60
 DB 252 GAGCGCGGCTCCCGGGCTGCTGTACAGAGATGGCAGGAGCGAGGAAATTACTTCC 311
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTrrpGlnGly 80
 DB 312 CCCAGTGTCCCAACAAGCCGAGCTGTCTGCTCCTGCTGAGCCGAGGCTGGAGGGC 371
 QY 81 TyrValSerAspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuLe 100
 DB 372 TAGTGAGGACATCAAGGCGCTTCTCAGTCATTTCAGACGACAGGAGGAGGCTCATC 431
 QY 101 GlnAlaAlaGlnGlnLeuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuLeuAla 120
 DB 432 CAGGCGCCGCCAGAGCTGTGTGTGATGACAGAGCCGACAGAGGAGAGGCTGCTGCT 491
 QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGlnThrArgAlaGluAspProPro 140
 DB 492 GACCTCTGCACAAAGTCAGCCAGAACATCGCGGCGAGACCGGCTGAGAGACCGCG 551
 QY 141 TrpPheGluGlyLeuGlnSerArgPheGlnSerLysSerGlyTyrLeuArgTyrSerCys 160
 DB 552 TGGTTTGAAGGCTTGGAGTCCGATTCAGAGCAAGTCTGGCTATCTGAGATACAGCTGT 611
 QY 161 GluSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
 DB 612 GAGAGCGGATCCGAGATTCCTGAGGAGGAGTGTGAGTCTTACCCCTCCACAGTGTGCG 671
 QY 181 GlnAlaGlnGluGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
 DB 672 GAGGCTCAGAGAGATTCCTGCGGGTCTGCTGCTCCTGAGTGCAGAGGCTCGGCTCATG 731
 QY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaAlaGlyLysSerArgLeuCysThrPro 220
 DB 732 CAGTACAAATGCGAGCTCTCGACAGAGAGGAGGAGGCGGCGCTGCTGACACGCG 791
 QY 221 GlnGlyTrrpPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 DB 792 GAGGCTGCTGCTCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 851
 QY 241 IleAsnProTyrSerAsnAlaGlyGlnSerArgIleLeuPheSerThrTrpAsnLeuAspHis 260
 DB 852 ATCAACCCCTACAGTACAGAGGAGAGAGATCTCTTACGACCTGGAACCTGGATCAC 911
 QY 261 IleTleGluLysLysArgThrIleIleProThrLeuValGluAlaIleLysGlnGlnAsp 280
 DB 912 ATATAGAAAAGAAAGAACGACCATCATCTCTACACTGTGAGAGAAATTAAGAAACAGAT 971
 QY 281 GlyArgGluValAspTrrpGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLys 300
 DB 972 GGAAGAGAAAGTGAAGTGGAGTATTTTATGGCTGCTTTTACCTCAGAGAACCTAAAA 1031
 QY 301 LeuValHisIleValCysHisLysLysThrHisLysLeuAsnCysAspProSerArg 320
 DB 1032 CTAGTGCACATTTGTCTGCTCAAGAAACCAACCAACCACTCACTGACCCGAGAGAGA 1091
 QY 321 IleTyrLysProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
 DB 1092 ATCTACAAACCCAGAACAGGTGAAGCGAGAGAGCTGTGGGGAAGACCCAG 1145
 RESULT 4
 AF039210 1017 bp mRNA linear PRI 17-APR-1998
 LOCUS Homo sapiens caspase-activated nuclease mRNA, complete cds
 DEFINITION AF039210
 ACCESSION AF039210
 VERSION AF039210.1 GI:3056726
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1017)

AUTHORS Halenbeck, R., Macdonald, H., Rouleston, A., Chen, T. T., Conroy, L. and Williams, L. T.
 TITLE CPAN, a human nuclease regulated by the caspase-sensitive inhibitor DEE45
 JOURNAL Curr. Biol. 8 (9), 537-540 (1998)
 MEDLINE 98228358
 PUBMED 9560346
 REFERENCE 2 (bases 1 to 1017)
 AUTHORS Halenbeck, R., Macdonald, H., Rouleston, A., Chen, T. T., Conroy, L. and Williams, L. T.
 TITLE Direct Submission
 JOURNAL Submitted (18-DEC-1997) Technologies R & D, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA
 FEATURES
 source
 1..1017
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="pancreas"
 1..1017
 /function="apoptotic endonuclease"
 /function="activation of caspases degrades inhibitor and activates CPAN to degrade chromatin"
 /note="CPAN: detected in cytosol in an inactive form bound to a caspase-sensitive inhibitor"
 /codon_start=1
 /product="caspase-activated nuclease"
 /protein_id="AAC39709.1"
 /db_xref="GI:3056727"
 /translation="MLQPKSVKRLALRSPKRGVAGRSCEYLIRKGLRFPOLPERS RLIEDGTELTEDYPPSPDMEALVLTITGAMOGVSPIRFLSFHFPQGLIA AQQLCDQAPORRLADLHNVSQIAETRAEDPWFEGLESFROSGLYRSC ESIRSLYREVSSYPTVGAQAEELFRLVLSQOKLRMQYNGSYFDRGAKGSRIC TPGWESGCPFMDSCLSRHSINPYNSRESRLIFSTWINDHIIIEKRTIIPLVLEAI KEDGREVDEYFYGLFTSENILKVIHIVCHKRTTKLNCDPSRIYKPYRLKRPV RKRO"
 BASE COUNT 236 a 293 c 303 g 185 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 3.65e-159 Length: 1017
 Score: 1786.00 Matches: 337
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.70% Mismatches: 0
 Query Match: 99.83% Indels: 0
 DB: 9 Gaps: 0
 US-09-748-451-2 (1-338) x AF039210 (1-1017)
 QY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
 DB 1 ATGCTCCAGAAAGCCCAAGAGCGGTGAMAGCTCGGGCCCTGGCGACGCCGAGGAAGTTCGGC 60
 QY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 DB 61 GTGGCTGGCGGAGCTCCAGAGAGTCTCGCGAAGGCTGTCTCCCTCCAGCTTCCCT 120
 QY 41 GluArgLysSerArgLeuCysLeuTyrGluAspGlyThrGluLeuThrGluAspTyrPhe 60
 DB 121 GAGGCGGCTCCCGGCTGCTGTACAGAGATGGCAGGAGCTGACGAGGAAGTACTTCC 180
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTrrpGlnGly 80
 DB 181 CCCAGTGTCCCAACAAGCCGAGCTGTCTGCTCCTGCTGAGCCGAGGCTGGAGGGC 240
 QY 81 TyrValSerAspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuLe 100
 DB 241 TAGTGAGGACATCAAGGCGCTTCTCAGTCATTTCAGACGACAGGAGGAGGCTCATC 300
 QY 101 GlnAlaAlaGlnGlnLeuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuLeuAla 120
 DB 301 CAGGCGCCGCCAGAGCTGTGTGATGACAGAGCCGACAGAGGAGAGGCTGCTGCT 360
 QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGlnThrArgAlaGluAspProPro 140

Db 361 GACCTCTCTCCACAACTGACCAACATGCGCGGACCAACCGGGGAGAGACCGCGC 420

QY 141 TTPPhGLUGLVLengLuserArpPhGLInSerLysSerGlyTYrLeuArpTYrSerCys 160

Db 421 TGGTTTGAAGGCTTGAGTCCCGCATTTCAAGCAAGATCGCATCTAGATACAGCGGT 480

QY 161 GLuserArpLleArpSerTYrLeuArpGLuValSerSerTYrProSerThnValGLyAla 180

Db 481 GAGAGCGGATCCGAGTTACTGAGGAGGTGAGTCTTACCCTCCACAGTGGGTGGC 540

QY 181 GLuValInGLUGLuserArpValLengLysSerMetCysGLuArpLeuArpSerMet 200

Db 541 GAGCTCAGAGAGATTCCTCGGGGTCTCGGCTCCATGTGCCAAGAGTCCGGTTCATG 600

QY 201 GLuTYrAenGLySerTYrPhenArpArgLYAlaLysGLyLysSerArpLeuCysThnPro 220

Db 601 CAGTACAAATGGACGACTCTTGACAGAGAGGACCAAGGGGAGCGCGCTCTGCACACCG 660

QY 221 GLUGLYTTPPhSerCysGInGLyProPhenArpMetAspSerCysLeuSerArpHisSer 240

Db 661 GAAGCTGGTCTCTCGGACAGGCGCTTTGACATGAGAGGTGGTTATCAAGACACTCC 720

QY 241 IleAsnProTYrSerAsnArpGLuSerArpLleLeuPhenSerThnTrpAsnLeuAspHis 260

Db 721 ATCAACCCCTACAGTAAAGGAGGAGAGCGAGATCGCTTGACACCTGGAGACCTGATCAC 780

QY 261 IleIleGLuLYLysArpThLleIleProThnLeuValGLuAlaIleLysGLuGLuAsp 280

Db 781 ATATATGAAAAGAAACGACCATCATCTTACACTGGTGAAGCAATTAAGAACAAAGAT 840

QY 281 GLYArpGLuValAspTYrGLuTYrPhenTYrGLyLeuLeuPhenSerGLuAsnLeuLys 300

Db 841 GGAAGGAAAGTGAGTGGAGTATTTTATGAGCTGCTCTTTACTCAGAGACTTAA 900

QY 301 LeuValHisLleValLysHisLysLysThnHisLysLeuAsnCysAspProSerArp 320

Db 901 CTAGTGCATATGTCTGCCATTAAGAAACCAACCCCAACTCACTGTGACCCAGCAGA 960

QY 321 IleTYrLysProGLuInThrArgLeuLysArpGLuInProValArgLysArgIn 338

Db 961 ATCTACAAACCCCGACACAGGTTTGAACGGAACAGCTGTGGGAAGCGCGAG 1014

RESULT 5
AB028913
LOCUS
AB028913
DEFINITION
Homo sapiens DFE40 delta mRNA for DNA fragmentation factor 40 kDa
AB028913
VERSION
AB028913.1 GI:13516478
KEYWORDS
DNA fragmentation factor 40 kDa subunit delta; DNA fragmentation factor 40 kDa subunit.
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (sites)
Nakagawara, A., Takahashi, M., Takada, N. and Kawamoto, T.
DFF40 delta
Published only in database (2001)
2 (bases 1 to 2926)
Nakagawara, A., Takahashi, M., Takada, N. and Kawamoto, T.
Direct Submission
Submitted (JUN-1999) to the GenBank database
Research Institute, Division of Biochemistry, 666-2 Niltona, Chugai, Chiba, Chiba 260-8717, Japan
(E-mail: takirane@chiba-cg.pref.chiba.jp,
tel:81-43-264-5431 (ex.3201), Fax:81-43-265-4459)
FEATURES
SOURCE
1..2926
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="Neuroblastoma"

PR1 03-APR-2001

gene	1. .2926
gene	/gene="DFF40 delta"
CDS	132. .443
	/gene="DFF40 delta"
	/product="DNA fragmentation factor 40 kda subunit delta"
	/protein_id="BAB40449.1"
	/db_xref="gi:13516479"
	/translation="MLOKPKSVKILRALRSPKRGVAGRSQGVLEKGLRQLEPERGS
	RLCIYEDGTELEDEDYFSPVSDNMAELVLTLLGQAMQMFCHVSQSDSLTLGSSCPALV
BASE COUNT	658 a 726 c 776 g 766 t
ORIGIN	S*
Alignment Scores:	
Score:	4,466-156
Score:	1759.50
Percent Similarity:	92.10%
Best Local Similarity:	91.83%
Query Match:	98.35%
Query Match:	9
Query Match:	1
US-09-748-451-2 (1-338) x AB028913 (1-2926)	
OY	1 MetLeuGlnLysProLysSerValLysLeuArqAlaLeuArqSerProArqLysPheGly 20
Db	132 ATGCTCCAGAACCCCAAGAGCGTGAAGCTGGGGCCCTGGCGACCCGAGAAAGTTGGCG 191
OY	21 ValAlaGlyArySerCysGlnGluValLeuArqLysGlyLysLeuArqPheGlnLeuPro 40
Db	192 GTGGCTGGCGGAGCTGGCCAGAGTCTGTGGCAAGGGCGTCTCCGCTCCAGCTCCCT 251
OY	41 GluAryGlySerAryLeuCysLeuTyrrGluAspGlyThrGluLeuThrGluAspTyrPhe 60
Db	252 GACCGGGGTCTCCGGCTGTGCTGTACAGAGATGGCACGAGCTGACGGAATATCTTC 311
OY	61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTrpGlnGly 80
Db	312 CCCAGTGTCTCCGACACGCGCAGAGCTGTGTCTCTACCTTGGCCAGCGCTGGACGGCC 371
OY	81 Tyr::: 81
Db	372 TGGTTTGGCCATGTTTCCAGAGATAGTCACACTCTGGGCTCAAGCTGTCTCTCCGCC 431
OY	82 ::::::::::::::::::::ValSerAspLLeuArqPheLeuSerAla 91
Db	432 TTGGTCTCTAAAGTCTGTGGGATTACAGATGTGTGGACATCTACAGCGCTTCTCTCA 491
OY	92 PheHisGluProGlnValGlyLeuLlGlnAlaAlaGlnLeuLeuGlnLysAspGlnGln 111
Db	492 TTTCACGACGCCACAGTGGGGCTCATCCAGGCCGCCACGACGCTGCTGTGTATGAGCAG 551
OY	112 AlaProGlnAryGlnAryLeuLeuAlaAspLeuLeuHisAsnValSerGlnAsnLLeuAla 131
Db	552 GCCCCACAAAGGCAAGAGCTCTGTGTACTCTCTGTGCACACATCGACCCAGAACATCTCCG 611
OY	132 AlaGluThrAryAlaGluAspProPTrPheGlnGlyLeuGlnLysAryPheGlnSer 151
Db	612 GCCGAAACCCGGGCTGAGAGACCCGCCCTGTGTATAGAGCTTGAAGTCCGGAATTTACAGCC 671
OY	152 LysSerGlyTyrLeuAryTyrSerCysGlnSerAryGlnLeuArySerTyrLeuAryGluAla 171
Db	672 AATCTGGGTATCTAGATACAGCTGTGAAGCCGAGTCCGGAGTAACTGTGAGGAGGTG 731
OY	172 SerSerTyrProSerThrValGlyAlaGluAlaGlnGluGluLysLeuAryValLeuGly 191
Db	732 AGCTCTACCCCTCCACAGTGGGTGGCGAGAGCTTCAGAGAAATTTCTGTGGGCTCTCCGCG 791
OY	192 SerMetCysGlnAryLeuArySerMetGlnTyrAsnGlySerLysPheAspAryGlyAla 211
Db	792 TCCATGTGCCAAGAGCTCCGGTTCATGACAGACAAATGGCAGCTACTTCTACACAGAGAACCC 851
OY	212 LysGlnGlySerAryLeuCysThrProGlnGlyTyrPheSerCysGlnGlyProPheAsp 231

Db 852 AAGGCGGACACCCCTCTGACACCGAAGGCTGTCCTCCAGGCTCCCTTAC 911
 QY 232 Metaspsercysleuserarqhiserlleasnprotyrseraanargluserarglle 251
 Db 912 ATGACAGACCTCTATCAAGACACTCCATCAACCCCTACACTAACAGGAGACGATC 971
 QY 252 Leupheserthrrpaasnleuaspshisilleglulysylsargthrlleleprothr 271
 Db 972 CTCTTACACACCTGGAACCTGATCACAATATAGAAAAGAACACCATCATCTCTACA 1031
 QY 272 Leuvalgluailllelyslglulnaspolyarglulvalaaptrpqlutyrpheyrcly 291
 Db 1032 CTGCTGGAAGCAATTAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1091
 QY 292 Leuuphehrsergluasnleuylsleuvalhisilevalcyshtlsylsargthrl 311
 Db 1092 CTGCTTTTACCTAGAGAACCTTAACACTGACATGCTGCTGCAATTAACAAAACCCAC 1151
 QY 312 Histylsleuasnrcysaspserarqllletrlysproglulnhrargleulysarglys 331
 Db 1152 CACAGAGCTCAACTGTGACCCGAGAGATCTACAAACCCAGACAGAGAGAGAGAGAGAG 1211
 QY 332 Gluprovalargylsargglu 338
 Db 1212 CAGCCTGTGCGGAACGCCAG 1232

too late

RESULT 6 2962 bp mRNA linear PRI 03-APR-2001
 AB028911 Homo sapiens DFF40 beta mRNA for DNA fragmentation factor 40 kDa
 LOCUS subunit beta, complete cds.
 DEFINITION
 ACCESSION AB028911.1 GI:13516474
 VERSION DNA fragmentation factor 40 kDa subunit beta.
 KEYWORDS Homo sapiens tissue_1lb:Fetal brain cdna to mRNA.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Nakagawara, A., Takahashi, M., Takada, N. and Kawamoto, T.
 TITLE DFF40 beta
 JOURNAL Published Only in Database (2001)
 REFERENCE 2 (bases 1 to 2962)
 AUTHORS Nakagawara, A., Takahashi, M., Takada, N. and Kawamoto, T.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUN-1999) Akira Nakagawara, Chiba Cancer Center
 Research Institute, Division of Biochemistry, 666-2 Nitona,
 Chuo-ku, Chiba, Chiba 260-8717, Japan
 (E-mail: akiranak@chiba-cs.pref.chiba.jp,
 Tel: 81-43-264-5431 (ex. 5201), Fax: 81-43-265-4459)
 FEATURES
 source location/Qualifiers
 1..2962
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_1lb="Fetal brain"
 1..2962
 /gene="DFF40 beta"
 132..944
 /gene="DFF40 beta"
 /codon_start=1
 /product="DNA fragmentation factor 40 kDa subunit beta"
 /protein_id="BAB40447.1"
 /db_xref="GI:13516475"
 /translation="MLOPKSVKRLALRSPKFGVAGRCOEVLKGLRQLPBGRS
 RLCLYEDTELEDEYPPSPVNAELVLTGOAMGVSDIRFLSAFHEQVGLIOA
 AOOICDEOAPORRLADLIANSONIAETRAEDPMPFGLSRRDSKSGYLARSC
 ESRIRSLREVSSTPSTVGAEAEFLRLVLSKQRLRSMQYNSYTDKRGKGSRLC
 TPGEWFCQCGFDMDSLSRHSINPYSNRESRIIFSTWMDHIDGVILCGP*"

Pred. No.: 5,63e-156 Length: 2962
 Score: 1758.50 Matches: 338
 Percent Similarity: 89.18% Conservative: 0
 Best Local Similarity: 89.18% Mismatches: 0
 Query Match: 98.30% Indels: 41
 DB: 9 Gaps: 1

US-09-748-451-2 (1-338) x AB028911 (1-2962)
 QY 1 MetleuglnlspolyserVallylsleuargalaleuargserProarglysphegly 20
 Db 132 ATGCTCCAGAAAGCCCAAGAGCTGAGCTGGGGCCCTGCCAGCCGAGAGATTGGCC 191
 QY 21 Valalaglyargsercysglngluvalleuargylslycysleuargpneglu 40
 Db 192 GTGGCTGCGGAGAGCTGCAGAGAGAGTGTGGCAAGGGCTGTCTCCGCTCCAGCTCCCT 251
 QY 41 Gluarqglyserarqleucysleutyrulaspolythrghluleuthrgluasptyrphe 60
 Db 252 GAGCGCGGTTCCCGCTGTGCTGTACAGAGATGGACAGAGCTGCAGGAAGATTACTTC 311
 QY 61 ProserValProaspasnagluleuvalleuaurhrleuglyglalatrpolngly 80
 Db 312 CCCAGTGTCCCGACACAGCCGAGCTGTGCTGCTCACTTGGCCAGGCTGCGCAGGCGC 371
 QY 81 Tyrvalseraspillearqargpneleuseralapherhlsigluoproglulnvalglyleuile 100
 Db 372 TATGTAGACGACATCAGCGCGCTTCCTCAGTCATTCACAGACCACAGGTGGGCTCATC 431
 QY 101 Glulalaglnglnleuileucysaspglulnlnarproglulnarqglulnleuula 120
 Db 432 CAGCGCGCCAGACACTCTGTGTATAGAGAGGCCCCACAGAGGACAGAGCTGTGCTT 491
 QY 121 AspleuenuhlsasnValsergluasnillealalagluhrarqalagluasprothro 140
 Db 492 GACCTCCGCAACAACGTGACGACAGATCCGCGGCGAGACCCGGCGTGAAGACCCCGCG 551
 QY 141 Trpnehluglyleuugluserarqphneglinserylserylyrrleuaurtyrsercys 160
 Db 552 TGGTTTGAAGCTGTGAGTCCCGATTTCAGAGCAAGTGTGCTTCTGATACAGCTGT 611
 QY 161 GluserarqlllearyserTyrleuarglulvaliserseTyrproserThValglyleuila 180
 Db 612 GAGAGCCGGATCCGAGTACCTGAGGAGGTGAGCTCTTACCCCTCCACAGTGGTCCG 671
 QY 181 GlualaglngluupheleuargValleuuglysermetcysglnarqleuargsermet 200
 Db 672 GAGGCTCGAGAGAAATTCCTCGGGTCTCGGCTCCAGTGCAGAGGCTCGGTCCATG 731
 QY 201 GlntyrasnnglyserTyrpneasparglyalalysgllyserarqleucysThrPro 220
 Db 732 CAGTACAAATGGCAGCTACTTTCAGACAGAGACCAAGGCGGCGGCTGTGCACACCG 791
 QY 221 GluglyTrrpnehrsercysglnglyProphaspmetaspsercysleuserarqhiser 240
 Db 792 GAGGCTGGTCTCTCTGCGAGGGTCTCTTGACATGAGCAGCTGTATACAAAGACTCC 851
 QY 241 Ileasnprotyrserasnargluserarqllleleupheserthrrpaasnleuaspshis 260
 Db 852 ATCAACCCCTACAGTAACAGGAGAGAGATCCCTTACAGACTGTGAACCTGGATCAC 911
 QY 261 Ile----- 261
 Db 912 ATAGATGAGATCTTGTCTGTGCGCCAGAGCTAAAGTGCAGTGGCAGATCATGCTCACT 971
 QY 261 ----- 261
 Db 972 GAAGCTGAACTCTCGGCTCAAGGAGATCTCTGCTCAGCTCTGAGCTCGCTGGGAC 1031
 QY 262 -----1leglulysylsargthrlleleprothrleuvalglualilleysgluln 279
 Db 1032 TACAGAAATAGAAAAGAAAGACCACTCATCTCTACACTGTGTGGAAGCAATTAAAGAACAA 1091

BASE COUNT 668 a 738 c 790 g 766 t
 ORIGIN
 Alignment Scores:

[illegible]

411 GUAAGGUGSERTATGAGCAGUAUUCUGUAASGCUYTHGCUUAUETHNGUUAAPYRPhE 60
125 GAGCGCGGTTCCCGCGTGGCTGTGACGAGATGCGACGAGACGTGACGGAAGTTATCTTC 184
OY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlyAlaIaTPGIngly 80
Db 185 CCAGATGTCCCAACAGCCGACGAGCTGGTCTCTACCTTGGCGAGGCTTGGCAGGGCC 244
OY 81 -Tyr-----ValSerAspIleArgArgPheLeuSerAlaPheHisGluProGlnVal 98
Db 245 TGTGGCAGAGTATGAGGAGATCAGAGCGCTTCTTCAGTGGATTTTACAGAGCCACAGGTGGG 304
OY 98 YLeuIleGlnAlaIaGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 118
Db 305 GCGTATCCAGACCGCCGACAGCTGCTGTGTGTATGACGAGCCGACAGAGGAGAGGCT 364
OY 118 ULeuAlaAspLeuLeuHisAsnValSerGlnAsnIleAlaIaGlnIaThrArgAlaGln 138
Db 365 GCTGGTGCAGCTCTCTCTCAACAGTCACGCAAGACATCGGGGCGAGACCCGGGCTGAGGA 424
OY 138 PProProThrPheGlnGlyLeuGlnUserArgPheGlnSerGlyssSerAlaIlyrLeuArgTy 158
Db 425 CCGCGCGGTTTGAAGGCTTGAGCTGCCATTTCCAGACAGAGCTGGCATGTGAGATA 484
OY 158 rSerCysGlnUserArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThya 178
Db 485 CAGCTGTGAGAGCGGATCCGAGTTACCTGAGGAGGAGTGAAGCTCTCAACCCCTCAAGST 544
OY 178 IGIYAlaGlnAlaGlnGlnGlnGlnPheLeuArgValLeuGlnGlySerMetCysGlnArgLeu 198
Db 545 GGGTGGCGAAGGCTCAGAGAGAAATTCCTCGGGGTCCCGGGCTCCATGTCCAGAGGCTCG 604
OY 198 gSerMetGlnTyrAsnGlySerTyrPheAspArgGlyAlaIaGlyGlySerArgLeuTy 218
Db 605 GTCCATGCACTCAAGAGGAGGACTTCTGCACAGAGAGGACGAGGCGGAGCCGACCTTG 664
OY 218 sThrProGlnGlyTyrPheSerCysGlnGlnIlyProPheAspMetAspSerCysLeuSer 238
Db 665 CACACCGGAAGGCTGTCTCTCTGCGAGGGTCCCTTTCAGACAGACAGCTGTATCAAG 724
OY 238 gHisSerIleAsnProTyrSerAsnArgGlnUserArgIleLeuPheSerThrTyrAsnLe 258
Db 725 ACACCCCATCAACCCCTACAGTACAGAGGAGAGACAGAGATCCCTTCAGACACCTGAACT 784
OY 258 uAspHisIleIleGlnUlysArgThrIleIleProThrLeuValGlnAlaIleLysG1 278
Db 785 GATCATCATATATGAAAAAAGAACGACCATCATCTCTACACTGGGAGAACCAATTAAAGA 844
OY 278 uGlnAspGlyArgGlnValaAspTyrGlnUlyrPheTyArgLeuLeuPheThrSerGln 298
Db 845 ACAGAGATGGAAGAGAGTGGAGTGGAGTATTTTATGGCTGCTTTTAACTCAAGAA 904
OY 298 nLeuUlysLeuValHisIleValIcysHisLysLysTrpThrHisLysLeuAsnCysAsp 318
Db 905 CCTAAACTGTGCACATGTGTGCATAAAGAAAACCAACCCACAGAGCTCAAGTGTAGCC 964
OY 318 oSerArgIleTyLysProGlnIaThrArgLeuLysArgLysGlnProValArgLysArgI 338
Db 965 AAGCAGATCTCAAAACCCACAGACAGAGTTGAAGCGGAACGAGCTGTGGGAAAGCCA 1024
OY 338 n 338
Db 1025 G 1025

RESULT 8
AF409061
LOCUS AF409061 1043 bp mRNA linear
DEFINITION Homo sapiens DNA fragmentation factor B truncated form II (DFFB)
ACCESSION AF409061
VERSION AF409061.1 GI:15553730
KEYWORDS Homo sapiens.

two late

FIG. 11-SEP-2000

ORGANISM Homo sapiens
 REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Bayasas,J.R., Yuste,V.J., Perera,R. and Comella,J.X.
 TITLE Characterization of splice variants of human caspase-activated Dnase that show CIDE-N structure and function
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1043)
 AUTHORS Comella,J.X., Bayasas,J.R., Yuste,V.J. and Perera,R.
 TITLE Direct Submission
 JOURNAL Submitted (13-AUG-2001) Ciencias Mediques Basiques, Universitat de Lleida, Av Rovira Route 44, Lleida 25198, Spain
 FEATURES
 source
 1..1043
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="IMR-5"
 /cell_type="neuroblastoma"
 1..1043
 /gene="DFFB"
 5..352
 /note="CIDE-N only protein; alternatively spliced"
 /product="DNA fragmentation factor B truncated form II"
 /protein_id="AA02006.1"
 /db_xref="GI:15553731"
 /translation="MLQPKSVKRLRSPKRGVAGSCOEVLKQCLFQLPERS
 RLCLYDGETLEDYEDYPPVSDNAELVLTGLGAMQCEGMGLMRCERHMLPQCISRA
 TCGAHGGRPAAY"
 BASE COUNT 242 a 296 c 314 g 191 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,52e-156 Length: 1043
 Score: 1754.50 Matches: 337
 Percent Similarity: 97.40% Conservative: 0
 Best Local Similarity: 97.40% Mismatches: 1
 Query Match: 98.07% Indels: 8
 DB: 9 Gaps: 1
 US-09-748-451-2 (1-338) x AF409061 (1-1043)
 QY 1 MetLeuGlnIlyProIySserValIySleuArGAlaIeuArGserProArGlySphGly 20
 DB 5 ATGCTCCAGAAAGCCCAAGAGCGTGAAGCTCGGGCCCTGGCAGCCCGAGAGATTGCG 64
 QY 21 ValAlaGlyArGserCySgInGluValIeuArGlySgIyCysIeuArGpHeGInIeuPro 40
 DB 65 GTGGCTGGCCGAGCGTCCAGAGAGTCTGCGCAGAGGCTGTCTCCGCTTCCAGCTCCCT 124
 QY 41 GluArGlySerArGleuCySleuArGlyIuArGpIyThrGluIeuThrGluAspTyrPhe 60
 DB 125 GAGCGGGGTCCCGGCTGTGCTGTACGAGAGATGGCAGGAGTGAAGATTACTTC 184
 QY 61 ProSerValIuArGpAspAlaGluIeuValIeuLeuThrIeuGlyGlnAlaTTPGInGly 80
 DB 185 CCCAGTGTCCCAACAAGCCGAGCTGTCTCTCACTTGGGCCAGGCTGGCAGGCG 244
 QY 81 -Tyr-----ValSerAspIleArGArGpHeuSerAlaPheH 93
 DB 245 TGTGAGTGGCAAGACTTTGGAGATGTGAGCGACATCAGCGCTTCTCAGTGATTTCA 304
 QY 93 SgIuProGInValGlyIeuIleGlnAlaAlaGInGInIeuLeuCySAspIuGlnAlaP 113
 DB 305 CGAGCCACAGGTGGGCTCATCCAGGCGCCGACGAGCTCTGTGATGATGACAGGCGCC 364
 QY 113 oGlnArGInArGleuAlaAspIeuAlaAspIeuHIsaSnValSerGInAsnIleAlaG 133
 DB 365 ACAGAGGACAGAGCTGTGCTGCTACCTCCGCAACAAGCTCAGCAACATCCGCGCGCA 424
 QY 133 uThrArGAlaGluAspProIotrPpHeGluGlyIeuGluSerArGpHeInSerIySse 153

DB 425 GACCCGGCTGAGGACCCGCCGTGTGAAGGCTTGAAGCTCCGATTTAGACCAAGTC 484
 QY 153 TGLTYLeuArGtyrSerCySgIuSerArGlyIeArSerTyIeuArGluValIeSse 173
 DB 485 TGGTATCTAGATACAGCTGTGAGACCGGATCCGAGATTACTGAGGAGGTGAGCTC 544
 QY 173 rTyProSerThrValIyAlaGluAlaGInGluIuPheuArGValIeuGlySerm 193
 DB 545 CTACCCCTCCACGCTGGGTGGGAGGCTCAGGAGGAATTCCTCGCGCTCCCTCCAT 604
 QY 193 tCyGluArIeuArGserIeGInTyranGlySerTyPheAspArgGlyAlaIySg 213
 DB 605 GTGCACAGGCTCCGCTCAGTACGACATGAGAGCTACTTGCACAGAGAGCAAGG 664
 QY 213 yGlySerArGleuCySthrProGluGlyTyIePheSerCySgInGlyIuProPheAspMetas 233
 DB 665 CGGAGCGGCTCGACACCGAGAGGCTGTCTCTCGCAAGGCTCCCTTGAATGCA 724
 QY 233 pSerCySleuSerArGHisSerIleAsnProTyIySerAsnArGluSerArGlyIe 253
 DB 725 CAGCTGCTTATCAAGACATCCATCAACCCCTACAGTACAGGAGAGAGATCCCTT 784
 QY 253 eSerThrTrpAsnIeuAspHisIleIleGluIySArGThrIleIleProhIeAr 273
 DB 785 CAGCAGCTGGAACTGGATCACAATATGAAAGAAAGCAACATTCCTACACTGGT 844
 QY 273 IGlualIleIySgInAspGlyArGluValAspTrpGluTyIeTyIeArGlyIe 293
 DB 845 GGAAGCAATTAAGAACACAGATGGAAGAAGATGAGATTTTATATGCGCT 904
 QY 293 uPheThrSerGluAsnIeuIySleuValHisIleValIySlySlyThrHisIy 313
 DB 905 TTTTACCTCAGAGAACTTAAACATGACATGCTGTCCATGAAAGAAACCCACCA 964
 QY 313 sIeuAsnCyAspProSerArGlyTyIySProGInThrArGleuArGlySgInP 333
 DB 965 GCTCACTGTGACCCAGCAAGATCTACAAACCCAGACAGATTGAAAGGAGAGAGCC 1024
 QY 333 oValArGlyArGIn 338
 DB 1025 TGTGCGGAAGCCGAG 1040
 RESULT 9
 AF409062 1071 bp mRNA linear
 LOCUS AF409062
 DEFINITION Homo sapiens DNA fragmentation factor B truncated form III (DFFB)
 ACCESSION AF409062
 VERSION AF409062.1 GI:15553732
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1071)
 AUTHORS Bayasas,J.R., Yuste,V.J., Perera,R. and Comella,J.X.
 TITLE Characterization of splice variants of human caspase-activated Dnase that show CIDE-N structure and function
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1071)
 AUTHORS Comella,J.X., Bayasas,J.R., Yuste,V.J. and Perera,R.
 TITLE Direct Submission
 JOURNAL Submitted (13-AUG-2001) Ciencias Mediques Basiques, Universitat de Lleida, Av Rovira Route 44, Lleida 25198, Spain
 FEATURES
 source
 1..1071
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /cell_line="IMR-5"
 /cell_type="neuroblastoma"
 1..1071
 /gene="DFFB"

CDS

5. 247
 /gene="DFRB"
 /note="CIDE-N only protein; alternatively spliced"
 /codon_start=1
 /product="DNA fragmentation factor B truncated form III"
 /protein_id="Val02007.1"
 /db_xref="GI:15553733"
 /translation="MLOKPKSVKLRALNSPKRGVAGRSCEVLRLKGLRFLPERGS
 RICLYEDGETELDEYFSPVDPNAELVLTLLIGAMOG"

BASE COUNT 242 a 306 c 319 g 204 t
 ORIGIN

Alignment Scores:

4.43e-155 Length: 1071
 1743.00 Matches: 337
 94.938 Conservative: 0
 94.938 Mismatches: 1
 97.438 Indels: 18
 Gaps: 1

US-09-748-451-2 (1-338) x AF409062 (1-1071)

OY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
 DB 5 ATGCTCCAGAACGCCAGAACGCGTACAGCTCGGAGCCCTGCGACGCCAGAAATTCGGC 64
 OY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 DB 65 GTGGCTGGCCGAGCGCCAGAGAGTCTCTCGCAAGGCGCTGCTCCGCTCCAGCTCCCT 124
 OY 41 GluArgGlySerArgLeuCysLeuArgLysPheGlnLysPheGlnLysPheGlnLysPhe 60
 DB 125 GAGCGCGGTTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 184
 OY 61 ProSerValProAspAsnAlaGlnLeuValLeuLeuThrLeuGlnAlaLysPheGlnLys 80
 DB 185 CCCAGAGTCCCGAACAGCGAGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
 OY 81 Tyr-----ValSer 83
 DB 245 TG-ATGAGCTTCTGCTGAGACCGGACCTTTGTTGCTCCATTGCTGCGAGATGTGAGC 303
 OY 84 AspLeuArgArgPheLeuSerAlaPheLeuGlnProGlnValGlyLeuLeuGlnAla 103
 DB 304 GACATCAGCGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 363
 OY 104 GlnGlnLeuLeuCysAspGlnGlnAlaPheGlnArgLysPheGlnLeuLeuAlaPheLeu 123
 DB 364 CAGCAGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 423
 OY 124 HisAsnValSerGlnAsnLysAlaAlaGlnThrArgAlaGlnLysProProThrPheGln 143
 DB 424 CACAAAGCTCAGCAGAACATCGGCGGCGGAGACCGGCTGAGACCGGCTGAGAGCCGG 483
 OY 144 GlyLeuGlnSerArgPheGlnSerLysSerGlyTyrLeuArgTyrSerCysGlnSerArg 163
 DB 484 GCGTTGGAGTCCGCAATTCAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
 OY 164 IleArgSerTyrLeuArgLysValSerSerTyrProSerThrValGlyAlaGlnLys 183
 DB 544 ATCCGAGATTCCTGAGGAGGTGAGCTCTACCCCTCCAGCGTGGTGGCGAGGCTGAG 603
 OY 184 GlnGlnPheLeuArgValLeuGlnLysSerMetCysGlnArgLeuArgSerMetGlnTyrAsn 203
 DB 604 GAGGAATTCCTGCGGGTCCCGCGCTCCATGTCAGAGAGCGCTCGGTCATGCAATGCAAT 663
 OY 204 GlySerTyrPheAspArgLysValAlaGlnLysGlySerArgLeuArgTyrProGlnLysTyr 223
 DB 664 GCGACGACTCTTCAGACAGAGAGCCAGGCGGCGGCGGCGGCTGCTGCAACCGAGAGCTGG 723
 OY 224 PheSerCysGlnLysPheAspMetAspSerCysLysLeuArgLysSerLysLeuAspPro 243
 DB 724 TTCCTGCGCGAGGCTCTTTCATGACAGCTGCTGCTTATCAGACACTCTCATCAACCC 783

OY 244 TyrSerAsnArgGlnSerArgLysLeuPheSerThrTyrPheAsnLeuAspHisLysLys 263
 DB 784 TACATGACAGGAGGAG 843
 OY 264 LysLysArgThrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 283
 DB 844 AGAAG 903
 OY 284 ValAspTyrGlyTyrPheTyrGlyLeuLeuPheThrSerGlnAsnLeuLysLeuHis 303
 DB 904 GTGAGCTGGAGATATTTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
 OY 304 IleValCysHisLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 323
 DB 964 ATGCTGCTGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1023
 OY 324 ProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
 DB 1024 CCCAGACAGAGGTGAG 1068

Too late

RESULT 10
 AB028912
 LOCUS
 DEFINITION Homo sapiens DFF40 gamma mRNA for DNA fragmentation factor 40 kDa subunit gamma, complete cds.
 ACCESSION AB028912.1 GI:13516476
 VERSION DFF40 gamma: DNA fragmentation factor 40 kDa subunit gamma.
 KEYWORDS Homo sapiens tissue; testis; testis cDNA to mRNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 REFERENCE 1 (bases) N. Takada, N. and Kawamoto, T.
 AUTHORS N. Takada, N. and Kawamoto, T.
 TITLE DFF40 gamma
 JOURNAL Published Only in Database (2001)
 REFERENCE 2 (bases 1 to 3008)
 AUTHORS N. Takada, N. and Kawamoto, T.
 TITLE DFF40 gamma
 JOURNAL Submitted (14-JUN-1999) Akira Nakagawa, Chiba Cancer Center Research Institute, Division of Biochemistry, 666-2 Mitono, Chuo-ku, Chiba, Chiba 260-8717, Japan
 E-mail: akirana@chiba-cc.pref.chiba.jp,
 Tel: 81-43-264-5431 (ex. 5201), Fax: 81-43-265-4459)
 location/Qualifiers

FEATURES
 source
 1. 3008
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /lssues="11b-testis"
 1. 3008
 /gene="DFF40 gamma"
 132. 482
 /gene="DFF40 gamma"
 /codon_start=1
 /product="DNA fragmentation factor 40 kDa subunit gamma"
 /protein_id="BAB0448.1"
 /db_xref="GI:13516476"
 /translation="MLOKPKSVKLRALNSPKRGVAGRSCEVLRLKGLRFLPERGS
 RICLYEDGETELDEYFSPVDPNAELVLTLLIGAMOGSVGRATRTDTSLSPEDCQ
 ALONGRCORLEL"

BASE COUNT 673 a 739 c 815 g 791 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.8e-153 Length: 3008
 Score: 1730.00 Matches: 337
 Percent Similarity: 85.32% Conservative: 0
 Best Local Similarity: 85.32% Mismatches: 1
 Query Match: 96.70% Indels: 57
 DB: 9 Gaps: 1

US-09-748-451-2 (1-338) x AB028912 (1-3008)

QY 1 MetLeuGlnIysProLysSerValIysLeuArgAlaLeuArgSerProArgLysPheGly 20
 DB 132 ATGCTCCAGAGAGCCCAAGAGAGGCTGAGAGGCGCCCTGGCGCAGCCGAGAGAAATGTCGCG 191
 QY 21 ValAlaGlyArgSerGlyGlnValLeuArgGlyGlyCysLeuArgPheGlnLeuPro 40
 DB 192 GTGCTGGCGGAGCTCCAGAGAGGCTGCGCAAGAGGCTGTCTCCGCTTCACCTCCCT 251
 QY 41 GlnArgLysSerArgLeuGlyCysLeuTyrGlnAspLysThrGluLeuThrGluAspTyrPhe 60
 DB 252 GAGGCGGCTTCGCGGCTGTGCTGTAGAGAGTGGACGAGCGTGGAGGAAGATTACTTC 311
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTyrGlnGly 80
 DB 312 CCCAGTGTCCCGCAACAGCCGAGCTGTGCTCTACCTTGCGCCAGCGCTGCGCAGGCG 371
 QY 81 -Tyr----- 81
 DB 372 TCTGTGTGTGTCAGGGCAGAGACAAAGACCCGGGACCTTCAGTCTGAGTCTGTGAT 431
 DB 81 ----- 81
 DB 432 TGGCAGGCCCTGGGGAATGGGGGAAGATGTGTGTCAGAGGCTCTTGTGACCGGGGCG 491
 QY 82 -----ValSerAs 84
 DB 492 GATGTGTCTTCTGTGACCGGACCTTTGTTGTTCCTCATGTGTGCGAGATGTGACGGA 551
 QY 84 PileArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuLeuGlnAlaGly 104
 DB 552 CATCAGGCGCTTCTCTGATGATTCAGAGAGCCAGAGAGGAGGCTCATCCAGGCGGCCA 611
 QY 104 ngInLeuLeuCysAspGlnAlaProGlnArgGlnArgLeuLeuAlaAspLeuLeuH 124
 DB 612 GCAGCTGCTGTGTGATGACAGAGGCCACAGAGGCAAGGCTGTGCTGACCTCCGCA 671
 QY 124 sAsnValSerGlnAsnIleAlaGluThrArgAlaGluAspProProTyrPheGluG 144
 DB 672 CAACGTCAGCCAGAACATCCGCGGAGACCCGCGGCTGAGAGACCCGCGGTGGTTGAAG 731
 QY 144 YLeuGlnSerArgPheGlnSerIysSerGlyTyrLeuArgTyrSerCysGluSerArg 164
 DB 732 CTGGAGTCCCGGATTTAGAGCAGAGCTGTGCTATCTAGATACAGCTGTGAGACCGGAT 791
 QY 164 eArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAlaGluAlaGln 184
 DB 792 CCGAGATTACCTGAGGAGGTGAGCTCTACCCCTCCACAGTGGGTGGGAGGCTCAGA 851
 QY 184 uGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsnG 204
 DB 852 GGAATTCCTCGGGGTCTCGGCTCCAGTCCAGAGAGCTCCGGTCCATGCAATGCAATGG 911
 QY 204 YSerTyrPheAspArgGlyAlaLysGlyLysArgLeuGlyThrProGluGlyTyrPhe 224
 DB 912 CAGCTACTTGGACAGAGAGCCAGAGGCGGCGGAGCCGCTCTGACACCGGAGGCTGCT 971
 QY 224 eSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSerIleAsnPro 244
 DB 972 CTCCTGCGAGGTCCTTTGACATGACAGCTGCTTATCAAGACATCCATCAACCCCTA 1031
 QY 244 rSerAsnArgGluSerArgIleLeuPheSerThrTyrAsnLeuAspHisIleIleGly 264
 DB 1032 CAGTACAGAGAGAGAGAGATCTCTTCACACCTGACACCTGATACATTAATAGAAA 1091
 QY 264 sLysArgThrIleIleProThrLeuValGluAlaIleLysGlnGlnAspArgLysVal 284
 DB 1092 GAAAGCGACCATATCTACCTAGCTGGAGAGCAATTAAGAAACAAATGAGAGAGT 1151
 QY 284 LasPTyGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLysLeuValHis 304
 DB 1152 GGACTGGAGATATTTTATGGCTGCTTTTACCTCAGAAACCTAAACTAGTGCAT 1211

QY 304 eValCysHisLysLysThrThrHisLysLeuAsnCysAspProSerArgIleTyrLysPr 324
 DB 1212 TGTCTGCCATTAAGAAACCAACCCAGCTCAACTGACCTGAGAGCAATCTCAAAACC 1271
 QY 324 OGInThrArgLeuLysArgLysGlnProValArgLysArgGln 338
 DB 1272 CCAGACAAAGTTGAAGCGGAAGAGAGCTGTGCGGAACGCCAG 1314
 RESULT 11
 E36548
 LOCUS 1038 bp DNA linear PAT 18-JUN-2001
 DEFINITION Caspase-activating DNase.
 ACCESSION E36548
 VERSION E36548.1 GI:13022711
 KEYWORDS JP 199239494-A/1.
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 1038)
 AUTHORS Shigekazu, N. and Masato, E.
 TITLE Caspase-activating DNase
 JOURNAL Patent: JP 199239494-A 1 07-SEP-1999;
 OSAKA BIO SCI KENKYUSHO
 COMMENT OS Homo sapiens (human)
 FN JP 199239494-A/1
 PD 07-SEP-1999
 PF 25-DEC-1998 JP 1998369093
 PR
 PI SHIGEKAZU MASATO, MASATO, ENRI
 PC C12N15/09, C07K16/40, C12N5/10, C12N9/22, C12N15/00, C12N5/00 CC
 FH Key Location/Qualifiers
 FT source 1..1038
 FT /organism="Homo sapiens (human)".
 FEATURES
 source 1..1038
 location/qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 236 a 293 c 307 g 202 t
 ORIGIN
 Alignment Scores:
 Pred. NO.: 2.95e-124 Length: 1038
 Score: 1416.00 Matches: 261
 Percent Similarity: 88.17% Conservative: 37
 Best local Similarity: 77.22% Mismatches: 38
 Query Match: 79.15% Indels: 2
 DB: Gaps: 1
 US-09-748-451-2 (1-338) x E36548 (1-1038)
 QY 1 MetLeuGlnIysProLysSerValIysLeuArgAlaLeuArgSerProArgLysPheGly 20
 DB 132 ATGCTCCAGAGAGCCCAAGAGAGGCTGAGAGGCGCCCTGGCGCAGCCGAGAGAAATGTCGCG 191
 QY 21 ValAlaGlyArgSerGlyGlnValLeuArgGlyGlyCysLeuArgPheGlnLeuPro 40
 DB 192 GTGCTGGCGGAGCTCCAGAGAGGCTGCGCAAGAGGCTGTCTCCGCTTCACCTCCCT 251
 QY 41 GlnArgLysSerArgLeuGlyCysLeuTyrGlnAspLysThrGluLeuThrGluAspTyrPhe 60
 DB 252 GAGGCGGCTTCGCGGCTGTGCTGTAGAGAGTGGACGAGCGTGGAGGAAGATTACTTC 311
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTyrGlnGly 80
 DB 312 CCCAGTGTCCCGCAACAGCCGAGCTGTGCTCTACCTTGCGCCAGCGCTGCGCAGGCG 371
 QY 81 -Tyr----- 81
 DB 372 TCTGTGTGTGTCAGGGCAGAGACAAAGACCCGGGACCTTCAGTCTGAGTCTGTGAT 431
 DB 81 ----- 81
 DB 432 TGGCAGGCCCTGGGGAATGGGGGAAGATGTGTGTCAGAGGCTCTTGTGACCGGGGCG 491
 QY 82 -----ValSerAs 84
 DB 492 GATGTGTCTTCTGTGACCGGACCTTTGTTGTTCCTCATGTGTGCGAGATGTGACGGA 551
 QY 84 PileArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuLeuGlnAlaGly 104
 DB 552 CATCAGGCGCTTCTCTGATGATTCAGAGAGCCAGAGAGGAGGCTCATCCAGGCGGCCA 611
 QY 104 ngInLeuLeuCysAspGlnAlaProGlnArgGlnArgLeuLeuAlaAspLeuLeuH 124
 DB 612 GCAGCTGCTGTGTGATGACAGAGGCCACAGAGGCAAGGCTGTGCTGACCTCCGCA 671
 QY 124 sAsnValSerGlnAsnIleAlaGluThrArgAlaGluAspProProTyrPheGluG 144
 DB 672 CAACGTCAGCCAGAACATCCGCGGAGACCCGCGGCTGAGAGACCCGCGGTGGTTGAAG 731
 QY 144 YLeuGlnSerArgPheGlnSerIysSerGlyTyrLeuArgTyrSerCysGluSerArg 164
 DB 732 CTGGAGTCCCGGATTTAGAGCAGAGCTGTGCTATCTAGATACAGCTGTGAGACCGGAT 791
 QY 164 eArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAlaGluAlaGln 184
 DB 792 CCGAGATTACCTGAGGAGGTGAGCTCTACCCCTCCACAGTGGGTGGGAGGCTCAGA 851
 QY 184 uGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsnG 204
 DB 852 GGAATTCCTCGGGGTCTCGGCTCCAGTCCAGAGAGCTCCGGTCCATGCAATGCAATGG 911
 QY 204 YSerTyrPheAspArgGlyAlaLysGlyLysArgLeuGlyThrProGluGlyTyrPhe 224
 DB 912 CAGCTACTTGGACAGAGAGCCAGAGGCGGCGGAGCCGCTCTGACACCGGAGGCTGCT 971
 QY 224 eSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSerIleAsnPro 244
 DB 972 CTCCTGCGAGGTCCTTTGACATGACAGCTGCTTATCAAGACATCCATCAACCCCTA 1031
 QY 244 rSerAsnArgGluSerArgIleLeuPheSerThrTyrAsnLeuAspHisIleIleGly 264
 DB 1032 CAGTACAGAGAGAGAGAGATCTCTTCACACCTGACACCTGATACATTAATAGAAA 1091
 QY 264 sLysArgThrIleIleProThrLeuValGluAlaIleLysGlnGlnAspArgLysVal 284
 DB 1092 GAAAGCGACCATATCTACCTAGCTGGAGAGCAATTAAGAAACAAATGAGAGAGT 1151
 QY 284 LasPTyGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLysLeuValHis 304
 DB 1152 GGACTGGAGATATTTTATGGCTGCTTTTACCTCAGAAACCTAAACTAGTGCAT 1211

REFERENCE 1 (bases 1 to 155585)
 AUTHORS Bray-Alten, S.
 TITLE Direct Submission
 JOURNAL Submitted (19-OCT-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 COMMENT requests: clonerequests@sanger.ac.uk
 On Oct 21, 2001 this sequence version replaced g1:15021313.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPeP; Information on the WormPeP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr9
 Rpl1-205M20 is from the library RPCR-11.1 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone Rpl1-205M20. It may be shorter because we sequence overlapping sections only once, except for a short overlap.
 The true right end of clone Rpl1-205M20 is at 155585 in this sequence. The true left end of clone Rpl1-55504 is at 84490 in this sequence. The true right end of clone Rpl1-334P12 is at 2000 in this sequence.

FEATURES
 source 1..155585
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /clone="Rpl1-205M20"
 /clone_id="RPCR-11.1"
 BASE COUNT 48210 a 32474 c 29896 g 45005 t
 ORIGIN

Alignment Scores:
 No: 6.62e-106 Length: 155585
 nt similarity: 1251.00 Matches: 268
 Local Similarity: 85.88% Conservative: 24
 Mismatches: 34
 Query Match: 69.93% Indels: 18
 Gaps: 3

US-09-748-451-2 (1-338) x AL353671 (1-155585)
 QY 1 MetleuGlnlyrProlyrServalysLeuArghAlaleuArghSerProArghlyrPhegly 20
 DB 105080 GTGCTCCAGAGCAGCAGACGACGCTGGGCGCTGTGACGCCAGAGAAATTGGC 105021
 QY 21 ValAlaGlyArghSerGlyGlnGluValLeuArghlyrGlyCysLeuArghPheGlnLeuPro 40
 DB 105020 ATGCGGGGAGAGAGCTGCGGAGAACTGCTGCTTGTACCTCCCT 104961
 QY 41 GluArghlyrSerArghLeuArghlyrGluAspGlyThrGluLeuThrGluAspTyPhe 60
 DB 104960 CG-----TACGAGATGGCATGAGTGCAGCGAGAGC----- 104930
 QY 61 ProSerValProAspAlaGluLeuValLeuLeuThrLeuGlyGlnAlaTrpGlnGly 80

DB 104929 -----GTCCTGACAGATCCGACCTGCTGCTCACCCTCAGCCAGCCGCTGCGACGGC 104876
 QY 81 TyrValSerAspIleArghArghPheLeuSerIlePheHisGlnPurGlnValGlyLeu 100
 DB 104875 TATGTAGTGCATATCGGGTCTGCTCTCAGTCTTTTGCCAGCCGACGCTGCTATC 104816
 QY 101 GlnAlaAlaGlnGlnLeuLeuCysAspGlnAla-ProGlnArghGlnArghLeuVal 120
 DB 104815 CAGCCCGCCGCGACGCTACTGTGTATGAGAGAGCCGCCACAGAGAGAGACTGTGGC 104756
 QY 120 AspLeuLeuHisAsnValSerGlnAsnIleAlaGluThrArghAlaGluAspProPr 140
 DB 104755 CGACCTCTGTACACGACGACCCAGACGTCGACCGACCGACCGGCTGAGACCCGCT 104696
 QY 140 OTrrPheGlnGlyLeuGlnSerArghPheGlnSerIleSerIleTyrrLeuArghTyrrSer 160
 DB 104695 GTGCTTTCAGAGCTTGGAGTCCCATTTTGGAAATAGTGTGTGTGATGATACACTG 104636
 QY 160 SgluSerArghIleArghSerTyrrLeuArghGluValSerSerTyrrProSerThrValGlyAl 180
 DB 104635 TGAAGCCGATCCGAGATTAAGTACAGAGAGTGAAGTCTGCTCATGATGGTGGTGC 104576
 QY 180 agluAlaGlnGlnGlnGlnPheLeuArghValLeuGlySerMetCysGlnArghLeuArghSer 200
 DB 104575 AGAGGCTCAGAGAGAAATTCGCGGGCTCTCCATGTGCCAGAAAGCTCAGTCTGT 104516
 QY 200 TglnTyrrAsnGlySerTyrrPheAspArghGlyAlaGlyGlySerArghLeuCysTyrrPr 220
 DB 104515 GCATGACGACGCGACGCTATAC-MACAGAGAGAGCCAGAGGTGGACGCTCTCTCAGACC 104457
 QY 220 OGluGlyTyrrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArghHis 240
 DB 104456 GGAAGGCTAGTTCCTCCGAGGCTCCCTTGATGACATGAGCGGCTGTGTAAATATCTC 104397
 QY 240 rilaAspProtyrrSerArghGlnSerArghIleLeuPheSerThrPheAsnLeuAspH 260
 DB 104396 CATGACCCCTACAGTAAAGAGAGAGAGAGAGGCTCTCTTCTGAGCTGGAACCTGATCA 104337
 QY 260 sileIleGlu-LysLysArghThrIleLeuProThrLeuValGlnAlaIleLysGlnGln 280
 DB 104336 CATGCTGGAACCAAAACCGACCATTCCTACAGTGTGAAGCATTAAGAACAGAG 104277
 QY 280 spGlyArghGluValAspTrpGluTyrrPheTyrrGlyLeuLeuPheThrSerGlnAsnLeu 300
 DB 104276 ATGGAAGAGAGAGTGGAGCTGGAAATATTTTGTGCTTCTTCACTCAGAGAACTAA 104217
 QY 300 ysluValAlaHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSer 320
 DB 104216 AATTAGTGT-----GTCGCCATTAAGAAACACCCACCAACCAAGCTATGACCTGAACA 104161
 QY 320 rglleTyrrLysProGlnThrArghLeuArghLysGlnProValArghLysGln 338
 DB 104160 GAATCTCAAAACCCAGACAGAGTGAAGAGAGAGTGGCTGTGAAGAGCCGAC 104105

RESULT 16
 LOCUS AL355811
 DEFINITION Homo sapiens chromosome 9 clone Rpl1-552A11 map p13.1-21.1, ***
 SEQUENCING IN PROGRESS ***; 26 unordered pieces.
 ACCESSION AL355811
 VERSION AL355811.4 GI:9863741
 KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 142970)
 AUTHORS McIlroy, K.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 requests: clonerequests@sanger.ac.uk

COMMENT

On Aug 21, 2000 this sequence version replaced g1.9213647.
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: BA562A11
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 18% of reads Chemistry:
 Dye-terminator Big Dye; 81% of reads
 Consensus quality: 127694 bases at least Q40
 Consensus quality: 133388 bases at least Q30
 Consensus quality: 136538 bases at least Q20
 Insert size: 140470; sum-of-contigs
 Insert size: 155298; 6.6% error; agarose-fp
 Quality coverage: 3.12% in Q20 bases; sum-of-contigs Quality
 coverage: 3.39% in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 26 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 4033: contig of 4033 bp in length
 4034 4133: gap of 100 bp
 4134 8168: contig of 4035 bp in length
 8169 8268: gap of 100 bp
 8269 14731: contig of 6463 bp in length
 14732 14831: gap of 100 bp
 14832 18803: contig of 3972 bp in length
 18804 18903: gap of 100 bp
 18904 21145: contig of 2242 bp in length
 21146 21245: gap of 100 bp
 21246 26247: contig of 5002 bp in length
 26248 26347: gap of 100 bp
 26348 30120: contig of 3773 bp in length
 30121 30220: gap of 100 bp
 30221 32561: contig of 2241 bp in length
 32562 32561: gap of 100 bp
 32562 39611: contig of 7050 bp in length
 39612 39711: gap of 100 bp
 39712 48337: contig of 8628 bp in length
 48338 48437: gap of 100 bp
 48438 50511: contig of 2074 bp in length
 50512 50611: gap of 100 bp
 50612 52986: contig of 2375 bp in length
 52987 53086: gap of 100 bp
 53087 59466: contig of 6380 bp in length
 59467 59566: gap of 100 bp
 59567 64909: contig of 5343 bp in length
 64910 65009: gap of 100 bp
 65010 69805: contig of 4796 bp in length
 69806 69905: gap of 100 bp
 69906 75879: contig of 5974 bp in length
 75880 75979: gap of 100 bp
 75980 81017: contig of 5038 bp in length
 81018 81117: gap of 100 bp
 81118 95015: contig of 13898 bp in length
 95016 95115: gap of 100 bp
 95116 99680: contig of 4565 bp in length
 99681 99780: gap of 100 bp
 99781 104130: contig of 4350 bp in length
 104131 104230: gap of 100 bp
 104231 110144: contig of 5914 bp in length
 110145 110244: gap of 100 bp
 110245 115096: contig of 4852 bp in length
 115097 115196: gap of 100 bp

FEATURES
source

* 115197 118257: contig of 3061 bp in length
 * 118258 118357: gap of 100 bp
 * 118358 126637: contig of 8280 bp in length
 * 126638 126737: gap of 100 bp
 * 126738 132827: contig of 6090 bp in length
 * 132828 132927: gap of 100 bp
 * 132928 142970: contig of 10043 bp in length.
 Location/Qualifiers
 1. 142970
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="9"
 /map="p13.1-21.1"
 /clone="RP11-562A11"
 /clone_id="RPCT11.2"
 1. 4033
 /note="assembly-fragment:01252"
 /vector="assembly-fragment:01252"
 /vector_end="SP6"
 /vector_side="left"
 4134. 8168
 /note="assembly-fragment:00004"
 /fragment_chain:1"
 8269. 14731
 /note="assembly-fragment:01032"
 /fragment_chain:1"
 14832. 18803
 /note="assembly-fragment:00497"
 /fragment_chain:2"
 18904. 21145
 /note="assembly-fragment:00407"
 /fragment_chain:2"
 21246. 26247
 /note="assembly-fragment:01246"
 /fragment_chain:3"
 26348. 30120
 /note="assembly-fragment:00027"
 /fragment_chain:3"
 30221. 32461
 /note="assembly-fragment:00019"
 /fragment_chain:3"
 32562. 39611
 /note="assembly-fragment:00023"
 /fragment_chain:3"
 39712. 48337
 /note="assembly-fragment:00055"
 /fragment_chain:3"
 48438. 50511
 /note="assembly-fragment:00090"
 /fragment_chain:3"
 50612. 52986
 /note="assembly-fragment:00183"
 /fragment_chain:3"
 53087. 59466
 /note="assembly-fragment:00202"
 /fragment_chain:3"
 59567. 64909
 /note="assembly-fragment:00332"
 /fragment_chain:3"
 65010. 69805
 /note="assembly-fragment:00342"
 /fragment_chain:3"
 69906. 75879
 /note="assembly-fragment:00400"
 /fragment_chain:3"
 75980. 81017
 /note="assembly-fragment:00476"
 /fragment_chain:3"
 81118. 95015
 /note="assembly-fragment:00586"
 /fragment_chain:3"
 95116. 99680
 /note="assembly-fragment:00653"
 /fragment_chain:3"
 99781. 104130
 /note="assembly-fragment:00679"
 /fragment_chain:3"
 104231. 110144
 /note="assembly-fragment:00746.0"
 /fragment_chain:3"
 110245. 115096
 /note="assembly-fragment:00950"
 /fragment_chain:3"
 115197. 118257
 /note="assembly-fragment:00973"
 /fragment_chain:3"
 118358. 126637
 /note="assembly-fragment:01164"
 /fragment_chain:3"
 126738. 132827
 /note="assembly-fragment:01263"
 /fragment_chain:3"

Db 760 CCATACGAGAACCGGACGACGAGAAATCCCTGTCACACCTGGAAATCTTGACCACCTGATA 819
 Oy 263 GUAUYSLSYARGTHRIIELEProThrlenuValGUAUAIIEtysGluGlnAspGlyArg 282
 Db 820 GAAAGAGAGAGAGATGTCATTCCTGCTCCGCCACAGCAGCAGGCAATAGAGCAAT 879
 Oy 283 GUAUVAAspTPGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLysLeuVal 302
 Db 880 GATATAAATGTCGATTTATTTTACAAACTCTGTTCCACACAGAGAAATCAACTGCTGTT 939
 Oy 303 HisLleValCysHISLSYLSYSThrThrlHisLysLeuAsnCysAspProSerArgLleYr 322
 Db 940 CACATCGCTGTCATTAAGAAAGGGCCCATGAGCTGAGCTGTGACTCCCGAATAATCTAC 999
 Oy 323 LysProGlnThrArgLeuLysArg 330
 Db 1000 AAA-----CGGCTCAAGAAC 1014

RESULT 19

AL691523 160716 bp DNA. linear HTG 12-AUG-2002
 Homo sapiens chromosome 1 clone RP13-531C17, *** SEQUENCING IN
 PROGRESS ***, in ordered pieces.

AL691523 5 GI:21955581
 HTG: HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 160716)

AUTHORS White,S.
 JOURNAL Direct Submission
 Submitted (08-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquy@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
 On Jul 25, 2002 this sequence version replaced gi:21732123.

COMMENT

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquy@sanger.ac.uk
 ----- Project Information
 Center project name: BBS31C17
 ----- Summary Statistics
 Assembly Program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 28 of reads
 Chemistry: Dye-terminator Big Dye; 97% of reads
 Consensus quality: 158827 bases at least Q40
 Consensus quality: 159096 bases at least Q30
 Consensus quality: 159180 bases at least Q20
 Insert size: 170984; 0.4% error; agarose-ef
 Quality coverage: 10.84x in Q20 bases; sum-of-contigs Quality
 coverage: 10.19x in Q20 bases; agarose-ef

FEATURES
 source
 1. 160716
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP13-531C17"
 /clone_1lb="RPCT-13.3"
 1. 160716
 /note="assembly_fragment:03871
 clone_end:SP6
 vector_side:left
 clone_end:T7
 vector_side:right"

misc_feature

1. 160716
 /note="assembly_fragment:03871
 clone_end:SP6
 vector_side:left
 clone_end:T7
 vector_side:right"

BASE COUNT 38997 a 41501 c 40997 g 39221 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.44e-26 Length: 160716
 Score: 410.00 Matches: 77
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 22.92% Indels: 0
 DB: 2 Gaps: 0

US-09-748-451-2 (1-338) x AL691523 (1-160716)

Oy 262 lIeGluYsLsYARGTHRIIELEProThrlenuValGUAUAIIEtysGluGlnAspGly 281

Db 144873 ATGAAAGAGAGAGATGTCATTCCTGCTCCGCCACAGCAGCAGGCAATAGAGCAAT 144932

Oy 282 ArgGluValAspTPGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLysLeu 301

Db 144933 AGAGAAAGTGGAGCTGGAGTATTTTATGCGCTGCTTTTACCTCAGAGAACCTAAACTA 144992

Oy 302 ValHisLleValCysHISLSYLSYSThrThrlHisLysLeuAsnCysAspProSerArgLle 321

Db 144993 GTGACATGTCGTCGATTAAGAAAGGGCCCATGAGCTGAGCTGTGACTCCCGAATAATCTAC 145052

Oy 322 TyrLysProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338

Db 145053 TACAACCCGACAGACAGGTTGACGCGAAGAGCGCTGTGCGGAACGCGAC 145103

RESULT 20

AK098413 1833 bp mRNA linear PRI 15-JUL-2002
 LOCUS AK098413/c

DEFINITION Homo sapiens CDNA FLJ25547 fls, clone JTH01487.
 ACCESSION AK098413
 VERSION AK098413.1 GI:21758419
 KEYWORDS oligo capping; fls (full insert sequence).
 SOURCE Homo sapiens thyroid cell_line:JCR CDNA to mRNA, clone_1lb:JTH
 Clone:JTH01487.

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS 1
 Niromiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
 Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
 Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
 Suzuki,Y., Hata,H., Nakagawa,K., Mizuno,S., Morinaga,M.,
 Kawamura,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H.,
 Nishikawa,T., Sugiyama,A., Kawakami,B., Nagai,K., Isogai,T. and
 Sugano,S.
 NEDO human cDNA sequencing project

JOURNAL Unpublished
 2 (bases 1 to 1833)
 SUGANO,S. and SUZUKI,Y.
 REFERENCE Direct Submission
 AUTHORS Submitted (08-JUL-2002) Sumio Sugano, Institute of Medical Science,
 University of Tokyo, Laboratory of Genome Structure, Human Genome
 Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
 (E-mail:cdna@elms.u-tokyo.ac.jp, Tel:81-3-5449-5286,
 Fax:81-3-5449-5416)

COMMENT

NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; CDNA full insert sequencing:
 Research Association for Biotechnology (RAB); CDNA library
 construction and 5'-end one pass sequencing: Institute of Medical
 Science, University of Tokyo, Laboratory of Genome Structure, Human
 Genome Center; 3'-end one pass sequencing: RAB; clone selection for
 full insert sequencing: RAB and Helix Research Institute.

FEATURES

source
 1. 1833
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="JTH01487"
 /cell_line="JCR"
 /tissue_type="thyroid"

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 15:21:08 ; Search time 121 Seconds
(without alignments)
277.028 Million cell updates/sec

Title: US-09-748-451-2
Perfect score: 1789
Sequence: 1 MLCRPSVRLALRSPKRCG.....SKTRPQRLKRPVAKKQ 338

Gapop 10.0, Gapext 0.5
375593 seqs, 9917265 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications, AA:*

- 1: /cgn2_6/piddata1/pubpaa/US08_NEM_PUB.pep.*
- 2: /cgn2_6/piddata1/pubpaa/PCU_NEM_PUB.pep.*
- 3: /cgn2_6/piddata1/pubpaa/US06_NEM_PUB.pep.*
- 4: /cgn2_6/piddata1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/piddata1/pubpaa/US07_NEM_PUB.pep.*
- 6: /cgn2_6/piddata1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/piddata1/pubpaa/PCUUS_PUBCOMB.pep.*
- 8: /cgn2_6/piddata1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/piddata1/pubpaa/US08_NEM_PUB.pep.*
- 10: /cgn2_6/piddata1/pubpaa/US08_PUBCOMB.pep.*
- 11: /cgn2_6/piddata1/pubpaa/US10_NEM_PUB.pep.*
- 12: /cgn2_6/piddata1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/piddata1/pubpaa/US08_NEM_PUB.pep.*
- 14: /cgn2_6/piddata1/pubpaa/US08_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	100.0	338	10 US-09-748-451-2	Sequence 2, Appl
2	114	6.4	76	9 US-10-165-800-13	Sequence 13, Appl
3	99	5.5	466	9 US-10-242-043-12	Sequence 12, Appl
4	99	5.5	466	12 US-10-152-647-4	Sequence 4, Appl
5	97	5.4	405	9 US-10-043-487-302	Sequence 302, Appl
6	97	5.4	465	9 US-09-935-642-3	Sequence 3, Appl
7	97	5.4	466	10 US-09-923-779-155	Sequence 155, Appl
8	97	5.4	466	12 US-10-152-647-4	Sequence 4, Appl
9	91	5.1	219	9 US-10-165-800-13	Sequence 13, Appl
10	89	5.0	331	10 US-09-748-451-2	Sequence 2, Appl
11	89	5.0	457	10 US-10-086-135-3	Sequence 3, Appl
12	87	4.9	366	10 US-09-923-779-155	Sequence 155, Appl
13	86.5	4.8	219	9 US-10-165-800-13	Sequence 13, Appl
14	86	4.8	219	9 US-10-165-800-13	Sequence 13, Appl
15	85.5	4.8	1458	9 US-10-203-824-3	Sequence 3, Appl
16	84.5	4.7	373	9 US-10-143-849-2	Sequence 2, Appl
17	83	4.6	2120	9 US-10-051-311A-2	Sequence 2, Appl
18	82.5	4.6	277	10 US-09-867-550-1316	Sequence 1316, Appl
19	82.5	4.6	633	9 US-10-097-340-19	Sequence 19, Appl

20	82.5	4.6	718	9 US-10-097-340-19	Sequence 19, Appl
21	82.5	4.6	751	9 US-10-097-340-18	Sequence 18, Appl
22	81.5	4.6	628	9 US-10-151-193-6	Sequence 6, Appl
23	81.5	4.6	630	9 US-10-151-193-5	Sequence 5, Appl
24	81.5	4.6	725	9 US-10-118-328-5	Sequence 5, Appl
25	81	4.5	1457	10 US-09-772-316-2	Sequence 2, Appl
26	80	4.5	480	9 US-09-859-888-5	Sequence 5, Appl
27	78.5	4.4	1618	9 US-09-963-875-1	Sequence 1, Appl
28	78.5	4.4	1618	9 US-10-136-891-2	Sequence 2, Appl
29	78.5	4.4	1618	9 US-10-136-891-2	Sequence 2, Appl
30	78	4.4	948	12 US-10-003-496-6	Sequence 6, Appl
31	78	4.4	775	9 US-08-820-843A-58	Sequence 58, Appl
32	78	4.4	1164	10 US-09-850-046A-2	Sequence 2, Appl
33	77.5	4.3	485	10 US-09-823-300-1664	Sequence 1664, Appl
34	77	4.3	382	9 US-10-187-666A-48	Sequence 48, Appl
35	77	4.3	382	9 US-09-987-107-37	Sequence 37, Appl
36	77	4.3	382	10 US-09-800-729-206	Sequence 206, Appl
37	77	4.3	629	9 US-10-197-666A-104	Sequence 104, Appl
38	77	4.3	629	10 US-09-799-875-17	Sequence 17, Appl
39	77	4.3	630	9 US-10-197-666A-102	Sequence 102, Appl
40	77	4.3	1331	10 US-09-801-368-370	Sequence 370, Appl
41	76.5	4.3	379	10 US-09-853-625B-11	Sequence 11, Appl
42	76.5	4.3	379	12 US-10-053-753-2	Sequence 2, Appl
43	76.5	4.3	679	9 US-10-278-547-42	Sequence 42, Appl
44	76.5	4.3	679	10 US-09-050-516-42	Sequence 42, Appl
45	76	4.2	1003	9 US-09-843-676-217	Sequence 217, Appl

ALIGNMENTS

RESULT 1	US-09-748-451-2	Application US/09748451
Sequence 2	Patent No. US20010011078A1	
GENERAL INFORMATION:		
APPLICANT: WANG, XIAODONG		
APPLICANT: LIU, XUESONG		
ATTORNEY: INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN APOPTOSIS		
FILE REFERENCE: US/09-748-451		
CURRENT FILING DATE: 2000-12-22		
PRIOR APPLICATION NUMBER: 09/061,702		
PRIOR FILING DATE: 1998-04-16		
NUMBER OF SEQ ID NOS: 21		
SOFTWARE: PatentIn Ver. 2.1		
SEQ ID NO 2		
LENGTH: 338		
TYPE: PRT		
ORGANISM: Homo sapiens		
US-09-748-451-2		
Query Match	100.0%; Score 1789; DB 10; Length 338;	
Best Local Similarity	100.0%; Pred. No. 8.3e-158;	
Matches 338; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MLCRPSVRLALRSPKRCG.....SKTRPQRLKRPVAKKQ 338	Sequence 19, Appl
DB	1 MLCRPSVRLALRSPKRCG.....SKTRPQRLKRPVAKKQ 338	Sequence 19, Appl
QY	61 PSYDPAELVLTICAGAGYSDIRFISAFHPRVGLQAQOILCDDEAPFORILTA 120	Sequence 120, Appl
DB	61 PSYDPAELVLTICAGAGYSDIRFISAFHPRVGLQAQOILCDDEAPFORILTA 120	Sequence 120, Appl
QY	121 DLHANTSONIAETRAEDPPPEFEGIESRFSKGYLRYSCESTRINSYREVSPSYGA 180	Sequence 180, Appl
DB	121 DLHANTSONIAETRAEDPPPEFEGIESRFSKGYLRYSCESTRINSYREVSPSYGA 180	Sequence 180, Appl
QY	181 EAGEEFLRVAGSMQCRRLSKMOYNSYFDRGAKGGSRLCTPEGWFSCQGFDMDCSLSHS 240	Sequence 240, Appl
DB	181 EAGEEFLRVAGSMQCRRLSKMOYNSYFDRGAKGGSRLCTPEGWFSCQGFDMDCSLSHS 240	Sequence 240, Appl
QY	241 INDYSNRESRLTSTWNLDTIEKRTIPLTVEAIKEDGREVMEYFYGLFTSENK 300	Sequence 300, Appl

Db 241 INPISNRSLRILFSTWIMLHTEKKRITIPPLVPAIKEDREVDWETFGILFSEMLK 300
QY 301 LVHIVCHKKTKHLNCDPSRIYKPPQTRLRKQPYRRKQ 338
Db 301 LVHIVCHKKTKHLNCDPSRIYKPPQTRLRKQPYRRKQ 338

RESULT 2

US-10-165-800-13
Sequence 13, Application US/10165800
Publication No. US20030092116A1
GENERAL INFORMATION:
APPLICANT: Chun, Myoung
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Kapeller-Libermann, Rosana
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: No. US20030092116A1 Nucleic Acid Sequences Encoding
TITLE OF INVENTION: Adenylate kinase, Phospholipid Scramblase-Like, DNA
TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,
TITLE OF INVENTION: and ATPase-Like Molecules and Uses Therefor
FILE REFERENCE: 35800/247838
CURRENT APPLICATION NUMBER: US/10/165,800
CURRENT FILING DATE: 2002-06-07

PRIOR APPLICATION NUMBER: 09/790,179
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,609
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,838
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/185,946
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,180
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,947
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/795,038
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/186,234
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/781,677
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/181,705
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13
LENGTH: 76
TYPE: PRT
ORGANISM: Artificial Sequence

OTHER INFORMATION: Consensus sequence for the CAD domain
-165-800-13

Query Match
Best Local Similarity 6.4%; Score 114; DB 9; Length 76;
Matches 30; Conservative 12; Mismatches 22; Indels 4; Gaps 3;

QY 14 RSPKFEVAGRSCEVLKGLRQPLPE--RGSRLCLYEDGTETL--TEDYFSPVDNAELV 70
Db 10 RNVAK--GVAASSTLELLSKVLDRKLPLDSEPLVLEEDGTVEDEDEFFTLNNELV 68
QY 71 LVLIGAM 78
Db 69 ALDEGEM 76

RESULT 3

US-10-242-943-12
Sequence 12, Application US/10242943
Publication No. US20030087412A1
GENERAL INFORMATION:
APPLICANT: Nandabalan, Krishnan

APPLICANT: Schulz, Vincent P.
APPLICANT: Yang, Melja
TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
FILE REFERENCE: 15966-521 NIK1 protein complexes
CURRENT APPLICATION NUMBER: US/10/242,943
CURRENT FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: US/09/167,206
PRIOR FILING DATE: 1998-10-06
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-10-242-943-12

Query Match
Best Local Similarity 5.5%; Score 99; DB 9; Length 466;
Matches 65; Conservative 47; Mismatches 11; Indels 90; Gaps 15;

QY 27 QEVLRKGLRQLPERSRLCLYEDGTETL--TEDYFSPVDNAELVLTIGAMGCVSDIR 86
Db 180 EDIMR---LRKLGEE---MLQREEAENTLQSFROVDNASTLARLDLKRKVESLQGEI 231
QY 87 RFLSAFHPQVGLQQA-----QQLCDQAPQORLLADLLHNVSONIAETRAEDPPW 141
Db 232 AFLKRLHEETIQELQAOIQEHQVQIDVDVSKPDLTALRDVROQY--ESVAKNLEQAEW 290
QY 142 F-----EGLESRFQSKGGLRY-----SCSRIR-----S 166
Db 291 YKSFADLSEANNNNDALRKAKQESTETFRQVQSLCEVDALKTGESLERQKREBN 350
QY 167 YLREVSSYPTVGAQAEFLRVLSMCQRLRSQYNGSYFDRGAKGSGRLCTPEGWFS 226
Db 351 FAVEAANYQDTIG--RLQDEIQNMKEAMRHLREYQ-----DLLNV 389
QY 227 QGPPDMSCLSRHSINPNSRSL-----FST-----WNLDI-----IEKRTIIP 272
Db 390 KMALDIEIATYRKLL---EGESRISLPLNFSSILNRETNLDSLPLVDTHSKRTL---L 443
QY 273 VEAIKEDDGRVD 285
Db 444 IKIVETRDGVIN 456

RESULT 4

US-10-152-647-4
Sequence 4, Application US/10152647
Patent No. US20020137110A1
GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. US20020137110A1uhhro,
APPLICANT: NAKANISHI, Keiko,
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 532.1004
CURRENT APPLICATION NUMBER: US/10/152,647
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US/09/610,401
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 466
TYPE: PRT
ORGANISM: Mus sp.
US-10-152-647-4

Query Match
Best Local Similarity 5.5%; Score 99; DB 12; Length 466;
Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;

QY 27 QEVLRKGLRQLPERSRLCLYEDGTETL--TEDYFSPVDNAELVLTIGAMGCVSDIR 86
Db 27 QEVLRKGLRQLPERSRLCLYEDGTETL--TEDYFSPVDNAELVLTIGAMGCVSDIR 86

Db 180 EDIMR--LREKIOE---MLOREESTLQSFROVDNANSLARLDLERRKVESLOEET- 231
QY 87 RFLSAFHEPOVGLIOA-----OQLLCDEQAPORRLADLLHNVSQNTAETRAEDPPW 141
Db 233 AFLKRLHEEIOELQAOIOEHQVIDVSKPDLTAALRDVROQY-ESVAANKIOEAEEM 290
QY 142 F-----EGLESRFQSKGYLR-----SCESRIR-----S 166
Db 291 YKSKFADLSEANNNMNLROAKQESNEYRQVQSLTCEVDALKTNESELEROMREMEEN 350
QY 167 YLREVSSTYVGAEOEFLRYLGSQCRLSQMGVNGSYFDRGAKGSRICPEGMFSC 226
Db 351 FALLEANQDITG-RLODEIQNNKEMARHLREYQ-----DLNV 389
QY 227 QGPFDMDSCLSRHSINYSNRESRI-----LFTWNLDHI-----IEKRTIIPPL 272
Db 390 KMALDIELATYRKL-----EGESRISLPLPTSSLNLEMTNLESIPLVDRHSKRTL---L 443
QY 273 VEAIKQDQREVD 285
Db 444 IKVETRDQVAIN 456

RESULT 5
US-10-043-487-302
; Sequence 302, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptic
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 302
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Shigella flexneri
US-10-043-487-302

Query Match 5.4%; Score 97; DB 9; Length 405;
Best Local Similarity 20.8%; Pred. No. 0.87; Indels 90; Gaps 15;
Matches 65; Conservative 46; Mismatches 112;
Db 27 QEYLKRGCLRFQLPERSRLCLYEDGTETEDYFSPVPAEVLVLTGQAMQGYSDIR 86
QY 119 EDIMR--LREKIOE---MLOREEAENTLQSFROVDNANSLARLDLERRKVESLOEET- 170
Db 87 RFLSAFHEPOVGLIOA-----OQLLCDEQAPORRLADLLHNVSQNTAETRAEDPPW 141
QY 171 AFLKRLHEEIOELQAOIOEHQVIDVSKPDLTAALRDVROQY-ESVAANKIOEAEEM 229
Db 142 F-----EGLESRFQSKGYLR-----SCESRIR-----S 166
QY 230 YKSKFADLSEANNNMNLROAKQESNEYRQVQSLTCEVDALKTNESELEROMREMEEN 289
QY 167 YLREVSSTYVGAEOEFLRYLGSQCRLSQMGVNGSYFDRGAKGSRICPEGMFSC 226
Db 290 FAVEAANKYDITG-RLODEIQNNKEMARHLREYQ-----DLNV 328
QY 227 QGPFDMDSCLSRHSINYSNRESRI-----LFTWNLDHI-----IEKRTIIPPL 272
Db 329 KMALDIELATYRKL-----EGESRISLPLPTSSLNLEMTNLESIPLVDRHSKRTL---L 382
QY 273 VEAIKQDQREVD 285
Db 383 IKVETRDQVAIN 395

RESULT 6
US-09-935-642-3
; Sequence 3, Application US/09935642
; Publication No. US20030044795A1
; GENERAL INFORMATION:
; APPLICANT: BYRJAISEN, Inger
; APPLICANT: LARSEN, Peter
; APPLICANT: STEPHEN, John
; TITLE OF INVENTION: Biochemical Markers for the Human
; FILE REFERENCE: 8969-014
; CURRENT APPLICATION NUMBER: US/09/935,642
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: PCT/GB97/02394
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: PCT/GB97/07132.8
; PRIOR FILING DATE: 1997-04-08
; PRIOR APPLICATION NUMBER: PCT/GB9618600.2
; PRIOR FILING DATE: 1996-09-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-935-642-3

Query Match 5.4%; Score 97; DB 9; Length 465;
Best Local Similarity 20.8%; Pred. No. 1.1; Indels 90; Gaps 15;
Matches 65; Conservative 46; Mismatches 112;
Db 27 QEYLKRGCLRFQLPERSRLCLYEDGTETEDYFSPVPAEVLVLTGQAMQGYSDIR 86
QY 179 EDIMR--LREKIOE---MLOREEAENTLQSFROVDNANSLARLDLERRKVESLOEET- 230
Db 87 RFLSAFHEPOVGLIOA-----OQLLCDEQAPORRLADLLHNVSQNTAETRAEDPPW 141
QY 231 AFLKRLHEEIOELQAOIOEHQVIDVSKPDLTAALRDVROQY-ESVAANKIOEAEEM 289
Db 142 F-----EGLESRFQSKGYLR-----SCESRIR-----S 166
QY 290 YKSKFADLSEANNNMNLROAKQESNEYRQVQSLTCEVDALKTNESELEROMREMEEN 349
QY 167 YLREVSSTYVGAEOEFLRYLGSQCRLSQMGVNGSYFDRGAKGSRICPEGMFSC 226
Db 350 FAVEAANKYDITG-RLODEIQNNKEMARHLREYQ-----DLNV 388
QY 227 QGPFDMDSCLSRHSINYSNRESRI-----LFTWNLDHI-----IEKRTIIPPL 272
Db 389 KMALDIELATYRKL-----EGESRISLPLPTSSLNLEMTNLESIPLVDRHSKRTL---L 442
QY 273 VEAIKQDQREVD 285
Db 443 IKVETRDQVAIN 455

RESULT 7
US-09-923-779-155
; Sequence 155, Application US/09923779
; Patent No. US20020076721A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Kelos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.553
; CURRENT APPLICATION NUMBER: US/09/923,779
; PRIOR FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155

LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-923-779-155

Query Match
Best Local Similarity 20.8%; Score 97; DB 10; Length 466;

Matches 65; Conservative 46; Mismatches 112; Indels 90; Gaps 15;

27 OEVLKRGCLRFOLPERGSRCLYEDGTELTEDYFPPVDPNAELVLLTGAMQGYSDIR 86
180 EDIMR---LREKLOEE---MLQREAEENTLOSFROVDNASTARLDLEKRVESLOEEI- 231
87 RFLSAFHEPOVGLIOA-----OQLCDEAPQORLLADLHNVSQNIATRAEDPPW 141
232 AFLKLIHEEIOELQIOEQHVQIDVDVSKPDLTALRDVROQY-ESVAKNLOEAEW 290
142 F-----EGLESRFQSKSGYLRY-----SCESRR-----S 166
291 YKSKFADLSEANRNDALQAKOESEYRQVOGLTCEVDALKGTNESLEROMREEN 350
167 YLREVSSTPYTGAEAQOEFLRVLGSQCRLRSQVNGSYFDGAKGSRCLTPEGWFS 226
351 FAVEAANYODTIG-RLODEIQNKKEEMARHLREYO-----DILNV 389
227 QGFPDMDCSLSRHSINYSRRESRL-----FST-----WNLDHI-----IEKRTIIP 272
390 KMALDIEIATYRKLL---EGEESRISLPLPFSLSLNRETNLDLPLVDYHSKRTE---L 443
273 VEAKEODGREVD 285
444 IKVETREDGOVIN 456

RESULT 8
US-10-152-647-3
Sequence 3, Application US/10152647
Patent No. US20020137110A1

GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. US20020137110A1uh1ro,
APPLICANT: NAKANISHI, Keiko,
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522.1004
CURRENT APPLICATION NUMBER: US/10/152,647
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: US/09/610,401
PRIOR FILING DATE: 2000-07-05
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-152-647-3

Query Match
Best Local Similarity 20.8%; Score 97; DB 12; Length 466;

Matches 65; Conservative 46; Mismatches 112; Indels 90; Gaps 15;

27 OEVLKRGCLRFOLPERGSRCLYEDGTELTEDYFPPVDPNAELVLLTGAMQGYSDIR 86
180 EDIMR---LREKLOEE---MLQREAEENTLOSFROVDNASTARLDLEKRVESLOEEI- 231
87 RFLSAFHEPOVGLIOA-----OQLCDEAPQORLLADLHNVSQNIATRAEDPPW 141
232 AFLKLIHEEIOELQIOEQHVQIDVDVSKPDLTALRDVROQY-ESVAKNLOEAEW 290
142 F-----EGLESRFQSKSGYLRY-----SCESRR-----S 166
291 YKSKFADLSEANRNDALQAKOESEYRQVOGLTCEVDALKGTNESLEROMREEN 350
167 YLREVSSTPYTGAEAQOEFLRVLGSQCRLRSQVNGSYFDGAKGSRCLTPEGWFS 226

351 FAVEAANYODTIG-RLODEIQNKKEEMARHLREYO-----DILNV 389
227 QGFPDMDCSLSRHSINYSRRESRL-----FST-----WNLDHI-----IEKRTIIP 272
390 KMALDIEIATYRKLL---EGEESRISLPLPFSLSLNRETNLDLPLVDYHSKRTE---L 443
273 VEAKEODGREVD 285
444 IKVETREDGOVIN 456

RESULT 9

US-10-165-800-14
Sequence 14, Application US/10165800
Publication No. US20030092116A1

GENERAL INFORMATION:

APPLICANT: Chun, Myoung

APPLICANT: Glucksmann, Maria Alexandra

APPLICANT: Kapeller-Libermann, Rosana

APPLICANT: Meyers, Rachel E.

TITLE OF INVENTION: No. US20030092116A1el Nucleic Acid Sequences Encoding

TITLE OF INVENTION: Adenylate Kinase, Phospholipid scramblase-Like, DNA

TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,

TITLE OF INVENTION: and ATPase-Like Molecules and Uses Therefor

FILE REFERENCE: 35800/247838

CURRENT APPLICATION NUMBER: US/10/165,800

PRIOR FILING DATE: 2002-06-07

PRIOR APPLICATION NUMBER: 09/790,179

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: 60/185,609

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: 09/790,838

PRIOR FILING DATE: 2001-02-22

PRIOR APPLICATION NUMBER: 60/185,946

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: 09/790,180

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 60/185,947

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: 09/795,038

PRIOR FILING DATE: 2001-02-26

PRIOR APPLICATION NUMBER: 60/186,234

PRIOR FILING DATE: 2000-02-29

PRIOR APPLICATION NUMBER: 09/781,677

PRIOR FILING DATE: 2001-02-21

PRIOR APPLICATION NUMBER: 60/181,705

PRIOR FILING DATE: 2000-02-10

NUMBER OF SEQ ID NOS: 25

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14

LENGTH: 219

TYPE: PRT

ORGANISM: Mus musculus

US-10-165-800-14

Query Match
Best Local Similarity 25.8%; Score 91; DB 9; Length 219;

Matches 46; Conservative 21; Mismatches 71; Indels 40; Gaps 7;

14 RSPRKFEVNAQRSCOEVLARKCLRFOLPERGSRCLYEDGTELTEDYFPPVDPNAELVLL 72
45 RYKAK-GTLAASIQEELDK-VLETLRLGVLTIVLEEDGTAVDSEDFOLLEDDTCIMVL 102
73 TLGQAW---OGYVS-----DIRFLSAFHEPO-----VGLIOAQQIL 107
103 ECGQSWSPKSGMLSTYGLGRKPKHSKDIAITTDYIKQNRDLFGSLNVAATYGYLGS 162
108 CDEAPQORLLADLHNVSQNIATRAEDPPWFGLESRFQSKSGYLRYSCESRRIR 165
163 CDQGVGPKRVVLELLRGTSSQL-----QGLGMLIGISSTLKHVYVAGADR 208

RESULT 10
 US-09-748-451-4
 Sequence 4, Application US/09748451
 Patent No. US20010011078A1
 GENERAL INFORMATION:
 APPLICANT: WANG, XINQONG
 APPLICANT: LIU, XUESONG
 TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN APOPTOSIS
 FILE REFERENCE: US/09/748, 451
 CURRENT APPLICATION NUMBER: US/09/748, 451
 CURRENT FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 09/061,702
 PRIOR FILING DATE: 1998-04-16
 NUMBER OF SEQ ID NOS: 21
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 4
 LENGTH: 331
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-748-451-4

Query Match 5.0%; Score 89; DB 10; Length 331;
 Best Local Similarity 22.2%; Pred. No. 3.7; Mismatches 146; Indels 78; Gaps 15;
 Matches 78; Conservative 49;

QY 5 PKSVKRLAIR-----SPKRGVAGRSQCEVLR-KGC--LRFQLPERSRLCYEDGT 53
 DB 10 PESGEIRTLKPCLLRRNYSRQHGVA--SLEDLRKACDILAIIDKSLTPVTLVAEDGT 68
 QY 54 EL-TEDYFSPVDPNAELVLTGLGAMQGVSDIRRLPSAFHEPOVGLIOAAQQLCDEQA 112
 DB 69 IYDDDYFLCLPSNKKFVALASNEKW-----AYNSDGTAMISOESFDVET 116
 QY 113 PORQLADLHNVSONIAETRAEDPPWFGLESFRQSGYLRYSCSRTSYLRVS 172
 DB 117 DSGAGL-----KKKNVARELK-----EDLSITLSEEDLQMLVDACSDLAQELR 162
 QY 173 SYPTGAEAGEEFLRVLSMCQRLRSQYNSYFDRGAKGSRLLCTPEGWFCGPPD- 231
 DB 163 OSCAFV-ORLQHTLQOVLDOREEVRSQKLLQLYLQALKEKESLSLKOE---ESKAARGE 218
 QY 232 ---MDSLSRHSINPYNSRESRLFTSMWLDHIIEKRTIIPVLEAKEDDREVDWE 287
 DB 219 EYDAVDTGMSRET-----SSDVALAS---HIL-----TALREKQAE----- 252
 QY 288 YFGILFTSENKIVHICHKTKTHKNCDSRLYKPPQRLKRPQVRKRO 338
 DB 253 ---LSLSQDLDELTKEDPKALVALNMDIKKTEYVQACERELALRLQO 299

RESULT 11
 US-10-086-135-3
 Sequence 3, Application US/10086135
 Publication No. US20020182677A1
 GENERAL INFORMATION:
 APPLICANT: Sheppard, Paul O.
 APPLICANT: Chandrasekhar, Yashin A.
 TITLE OF INVENTION: PANCREATIC AND OVARIAN POLYPEPTIDE
 FILE REFERENCE: 98-24
 CURRENT APPLICATION NUMBER: US/10/086,135
 CURRENT FILING DATE: 2002-02-26
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/366,448
 PRIOR FILING DATE: EARLIER FILING DATE: 1999-08-03
 PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,199
 NUMBER OF SEQ ID NOS: 35
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 3
 LENGTH: 457
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-10-086-135-3

Query Match 5.0%; Score 89; DB 9; Length 457;
 Best Local Similarity 21.8%; Pred. No. 5.7; Mismatches 104; Indels 104; Gaps 16;
 Matches 60; Conservative 40;

QY 44 SRLCYEDGTETEDYFSPVDPNAELVLTGLGAMQGVSDIRRLPSAFHEPO--VGLI 100
 DB 9 SLVLTVMGTELTQ-----VLPNPE-----ESMOY-----SSAQDSGRCITV 49
 QY 101 QAAQQLCDEQAPORO-RLADLHNVSONIAETRAE--AEDPPWFE-----GLESR- 149
 DB 50 VAPQGMCSRDARTKQRLRLEKYNMSQSIYVDRRTQRLQYKEMENQKGLSEKFR 109
 QY 150 -----OSKSGYL-----YSCSR-IRSYLREVSYPST----- 177
 DB 110 QVESHKQHLARQFAIKAKMDLRLPLVLEEKADAKVLOFKEEVQNLTSVLNLEQ 169
 QY 178 -VGAEGAEFEFLRVLSMCQRLR-----SMQYNSYFDRGAKGSR- 217
 DB 170 EIGAVDYDELQSRVSNLEERLRACQKLAGCKLIGSDPYTVKTSGR-----GSMNT 223
 QY 218 --GTPGS---WFCGCPDMDSCLSRHSINPYNS 246
 DB 224 DPLAREQDNRVWY-----MDGYHNRFTVEYKNS 251

RESULT 12
 US-09-925-300-1495
 Sequence 1495, Application US/09925300
 Patent No. US20020151681A1
 GENERAL INFORMATION:
 APPLICANT: Craig Rosen,
 APPLICANT: Steve Ruben,
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA101
 CURRENT APPLICATION NUMBER: US/09/925,300
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05988
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 1890
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 1495
 LENGTH: 366
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-925-300-1495

Query Match 4.9%; Score 87; DB 10; Length 366;
 Best Local Similarity 22.4%; Pred. No. 6.4; Mismatches 131; Indels 100; Gaps 18;
 Matches 81; Conservative 50;

QY 5 PKSVKRLAIR-----SPKRGVAGRSQCEVLR-KGC--LRFQLPERSRLCYEDGT 53
 DB 45 PESGEIRTLKPCLLRRNYSRQHGVA--SLEDLRKACDILAIIDKSLTPVTLVAEDGT 103
 QY 54 EL-TEDYFSPVDPNAELVLTGLGAMQGVSDIRRLPSAFHEPOVGLIOAAQQLCDEQA 108
 DB 104 IYDDDYFLCLPSNKKFVALASNEKW-----AYNSDGTAMISOESFDVET 151
 QY 109 DEQAPQRLADLHNVSONIAETRAE-----DPPWFGLESFRQSGYLRYSCS- 162
 DB 152 DSGAGLKKWVAKQLKEDLSITLSEEDLQMLVDACSDLAQELR-----LQSCAFV 203
 QY 163 -RISYLRVSSYPTVGAEGAEFLRVLSMCQRLRSQYNSYFDRGAKGSRLLCTPE 221
 DB 204 QRLQHTLQOVLDOREEVRSQKLLQLYLQALKEKESLSLKOE---ESKAARGE 245
 QY 222 GWFSCGPPD---MDSLSRHSINPYNSRESRLFTSMWLDHIIEKRTIIPVLEAK 276
 DB 246 ---ESKAARGEYDAVDGTISRET-----SSDVALAS---HIL-----TAL 280

QY 277 KEOGKREVDWETFFGLFTSENKLVHIVCHKTTHKINCPSRIYKQOTLRKKRPYK 336
Db 281 REKQAPF-----LSLSODELWYKEDPKALAVALNMDIKRTETVQACEERLRL 332
QY 337 RQ 338
Db 333 QQ 334

RESULT 13

US-10-165-800-11
Sequence 11, Application US/10165800
Publication No. US20030092116A1

GENERAL INFORMATION:

APPLICANT: Chun, Miyoung
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Kapeller-Liebermann, Rosana
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: No. US20030092116A1el Nucleic Acid Sequences Encoding
TITLE OF INVENTION: Adenylate Kinase, Phospholipid Scramblase-Like, DNA
TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,
FILE REFERENCE: 35800/247838
CURRENT APPLICATION NUMBER: US/10/165,800

PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 09/790,179
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,609
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,838
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/185,946
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,180
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,947
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/795,038
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/186,234
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/781,677
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/181,705
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-10-165-800-11

Query Match
Best Local Similarity 4.8%; Score 86.5; DB 9; Length 219;
Matches 47; Conservative 25; Mismatches 68; Indels 49; Gaps 8;

QY 14 RSPKRGVAGRSCEVLARKGLRQLPERSRLCLYEDGTEL-TEDYFSPVDNAELVLL 72
Db 45 RTIRK-GILTAATROELLAK-ALETLLNGVLTVLVEEDGTAVDESDFOQLLEDDTCLAVL 102
QY 73 TLGQAW-----QGIYS-----DIRFSLAFHEPQ-----VGLIQAOQL 106
Db 103 QSGGSWSPTRSGVLSYGIGRERPRKSRDIAFTFDVYKQNRDLFGSLNVAATFGYGLYSM 162
QY 107 LLDQAOAORORLLADLLHNVSQNTAAETRADPPMFEGLSERFOSKSGYLVSCE----- 161
Db 163 SCDFQGGGPKKVLBELRLWTSTLL-----QGLGMILGISITLNHAYGAQOW 210
QY 162 ---SRINSY 167
Db 211 QOKGRHSY 219

RESULT 14

US-10-165-800-15
Sequence 15, Application US/10165800
Publication No. US20030092116A1

GENERAL INFORMATION:

APPLICANT: Chun, Miyoung
APPLICANT: Glucksmann, Maria Alexandra
APPLICANT: Kapeller-Liebermann, Rosana
APPLICANT: Meyers, Rachel E.
TITLE OF INVENTION: No. US20030092116A1el Nucleic Acid Sequences Encoding
TITLE OF INVENTION: Adenylate Kinase, Phospholipid Scramblase-Like, DNA
TITLE OF INVENTION: Fragmentation Factor-Like, Phosphatidylserine Synthase-Like,
FILE REFERENCE: 35800/247838
CURRENT APPLICATION NUMBER: US/10/165,800

PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 09/790,179
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,609
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,838
PRIOR FILING DATE: 2001-02-22
PRIOR APPLICATION NUMBER: 60/185,946
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/790,180
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/185,947
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/795,038
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: 60/186,234
PRIOR FILING DATE: 2000-02-29
PRIOR APPLICATION NUMBER: 09/781,677
PRIOR FILING DATE: 2001-02-21
PRIOR APPLICATION NUMBER: 60/181,705
PRIOR FILING DATE: 2000-02-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15
LENGTH: 219
TYPE: PRT
ORGANISM: Homo sapiens
US-10-165-800-15

Query Match
Best Local Similarity 4.8%; Score 86; DB 9; Length 219;
Matches 24; Conservative 11; Mismatches 29; Indels 2; Gaps 2;

QY 14 RSPKRGVAGRSCEVLARKGLRQLPERSRLCLYEDGTEL-TEDYFSPVDNAELVLL 72
Db 44 RSSRR-GVMASLSQELISKTDALVATGLVTLVEEDGTVADTEEFQOTIGDWTWFML 102
QY 73 TLGQAW 78
Db 103 EKGQKW 108

RESULT 15

US-10-203-224-3
Sequence 3, Application US/10203224
Publication No. US20030086945A1

GENERAL INFORMATION:

APPLICANT: COLLINS, James E.
APPLICANT: FAABERG, Kay S.
TITLE OF INVENTION: PROLINE REPRODUCTIVE AND RESPIRATORY SYNDROME VIRUS AND
FILE REFERENCE: 110,01250101
CURRENT APPLICATION NUMBER: US/10/203,224
CURRENT FILING DATE: 2002-08-07
PRIOR APPLICATION NUMBER: PCY/US01/04351
PRIOR FILING DATE: 2001-02-08

PRIOR APPLICATION NUMBER: 60/181,041
PRIOR FILING DATE: 2000-02-08
PRIOR APPLICATION NUMBER: 60/193,220
PRIOR FILING DATE: 2000-03-30
PRIOR APPLICATION NUMBER: 60/206,624
PRIOR FILING DATE: 2000-05-24
PRIOR APPLICATION NUMBER: 60/215,373
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: 60/260,041
PRIOR FILING DATE: 2001-01-05
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patent Ver. 2.1
SEQ ID NO: 3
LENGTH: 1458
ORGANISM: Porcine reproductive and respiratory syndrome virus
0-203-224-3
TYPE: PR1
Query Match 4.8%; Score 85.5; DB 9; Length 1458;
Best Local Similarity 22.1%; Pred. No. 59;
Matches 64; Conservative 36; Mismatches 96; Indels 93; Gaps 15;
QY 100 IOAQQQLCDEQAPQORLLADLHNSQNTAETRA-----EDPPPEGL 145
DB 553 IEAGRLV-----PNDRLALALAHMAONASSETYASAAIIMDSACACIDHDPEMTIDL 607
QY 146 ---ESRFQSKGY-----LRYSCSRIRSYLREVSYPSTVGAQAQEEFLRLVIG-SMCO 195
DB 608 ICGIAOQARQDGYSPFPAPFMSWNERLRSH-NEGKKFRHCGICDAKADYASACGLDLC- 665
QY 196 RLRSQYNGSYEDR-----GANGSRLCTPBGWFSGCGPF-----DMSCLSRHSIN 242
DB 666 -----LFHSHFHQHPVYLSCGHAGSRECS-----QCOSPGAGRSPLDVLKQIPYK 714
QY 243 PYSNRESRILEFTWIND-----HIEKRTIIPVLVAIKEDGREVDE---YFY 290
DB 715 PPRVIVKRVGNKTTALDPRGYSRRLVAVKRGIA-----GNEVDLSDDGYGV 763
QY 291 GLPFTSENTLKVHIVCH-----KKTTHKLN--CDPSRIYKP 324
DB 764 PLPPTCKDIMWVACNVLLSKFIYPPGSGKTTWLLGOVDDDVITYTP 812
RESULT 16
US-10-143-849-2
Sequence 2, Application US/0143849
Publication No. US20030074699A1
GENERAL INFORMATION:
APPLICANT: Coupland, George M.
Pulterill, Joanna J.
TITLE OF INVENTION: Genetic control of flowering
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderyhe PC
STREET: 8th Floor, 1100 NO. US20030074699A1th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/143,849
FILING DATE: 14-May-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/516,191
FILING DATE: <unknown>
APPLICATION NUMBER: 08/945,056
FILING DATE: <unknown>

APPLICATION NUMBER: GB 9422083.7
FILING DATE: 02-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Ms. Mary J. Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-17
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-143-849-2
Query Match 4.7%; Score 84.5; DB 9; Length 373;
Best Local Similarity 22.6%; Pred. No. 11;
Matches 47; Conservative 26; Mismatches 64; Indels 71; Gaps 10;
QY 153 SGYLRSCSRIRSYLREVSYPSTVGAQAQEEFLRLVIGSMCORLSMQYNGSYFDRGAK 212
DB 36 SAYLCMSQDAQVHSANRVAS-----RHKRVYCS-CERAP- 71
QY 213 GGSRLCTPBGWFSGCGPFDMDSCLR-HSINPYSNRESRI-----LEFTWINDHIE 263
DB 72 -ATLCADADASIC-----TACDEVHSANPLARRQVPLIPISGSSMTTHOS 123
QY 264 KRTIIPVLVAIKEDGREVDE---WEY-----FYGLFTSENTKIV- 302
DB 124 EKMTDEKRLVVDGSESGDKAKEVASLFPNSKNNNNNNGLFSDRYLNDVYN 183
QY 303 ---HIVCAKKTTHKLNCDPSRIYKQPT 326
DB 184 SMDKFTGEGYSQHQQNS-----VPQT 206
RESULT 17
US-10-051-311A-2
Sequence 2, Application US/10051311A
Patent No. US20020164620A1
GENERAL INFORMATION:
APPLICANT: Boehringer Ingelheim International GmbH
TITLE OF INVENTION: Method for identifying compounds that modulate sister
FILE REFERENCE: 0652, 2290001
CURRENT APPLICATION NUMBER: US/10/051,311A
PRIOR FILING DATE: 2002-06-20
PRIOR APPLICATION NUMBER: EP 01 101 252.3
PRIOR FILING DATE: 2001-01-19
PRIOR APPLICATION NUMBER: US 60/297,440
NUMBER OF SEQ ID NOS: 13
SOFTWARE: Patent Ver. 2.1
SEQ ID NO: 2
LENGTH: 2120
TYPE: PR1
ORGANISM: Homo sapiens
US-10-051-311A-2
Query Match 4.6%; Score 83; DB 9; Length 2120;
Best Local Similarity 18.8%; Pred. No. 1.7e+02;
Matches 66; Conservative 52; Mismatches 130; Indels 104; Gaps 14;
QY 11 RALRSRPFKFGVNGSCOE-----VLKKGCLRPQLPERGS-----RLCYEDGT 53
DB 197 RAVAAHQIFDASGHGNEADQFDLIDLSRHVIRALVGERSSGGLISPORALCILE--- 253
QY 54 ELTEDYFSPVDNAELVLTLLGOAGQGVSDIRRL-SAFHEPQVGLIOAQQQL----- 106
DB 254 -----LTL-----EHCRFRQWRSHHDKAISAVEKASLYLHNTWL 287

QY 107 -----LCD-----EQAPORRLADLIHNSONIAETRAEDPPWEGLESROSK 152
 DB 288 APSLQOLQVGLKVLQVEBQ---AAVAKLIRKASAVLSKSMKSPSP---LALALYESC 340
 QY 153 SGYLRCSESRIRSEYLRVSSYSTVGAEQEEFLRVLGSMQCLRSQVNGSYFEDGAK 212
 DB 341 QFLLSGLEGRKTRK---RDLALSLFAPLGGYCSLQOQRDGYT---GGS 386
 QY 213 GGSRLCTPEGWFCQGFPMDSCLSRHSINPYSNRESRIIFSFWNLHIEKKRTIIPYL 272
 DB 387 SKQOQSFLOMYF--QG-----LHLYTVVYDFACQOQIVDLADLQGLVDSKSTVWM 437
 QY 273 VEAKEDQGRVMEYRYGGLFTSENKLVIYCHKTKTKLNDPSPRIYF 324
 DB 438 LEALEGLSGOELDHDGMGTASYTSN-----LAYSFYSRLVAFACALSEP 482

RESULT 18

US-09-867-550-1316

Sequence 1316, Application US/09867550

Genet No. US20020082206A1

GENERAL INFORMATION:

APPLICANT: Leach, Martin D.

APPLICANT: Mehrahan, Fuad,

APPLICANT: Conley, Pamela

APPLICANT: Law, Debbie

APPLICANT: Topper, James

TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and

FILE REFERENCE: 21402-013 (Cura-313)

CURRENT APPLICATION NUMBER: US/09/867, 550

CURRENT FILING DATE: 2001-09-20

PRIOR APPLICATION NUMBER: USN 60/208,427

PRIOR FILING DATE: 2000-05-30

NUMBER OF SEQ ID NOS: 2125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1316

LENGTH: 277

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: VARIANT

LOCATION: (1)

OTHER INFORMATION: wherein Xaa may be any one of Leu or Met or Val

Query Match 4.6%; Score 82.5; DB 10; Length 277;
 Best Local Similarity 23.8%; Pred. No. 11;
 Matches 45; Conservative 20; Mismatches 51; Indels 73; Gaps 8;

QY 165 RSYLREYSSYSTVGAE-----QEFPLVIGSMQCLRSQVNGSYFEDGAKGSR 217
 DB 20 RRCIRERAAAMAAALAAVAVASRLERQEDIRFWLSEVERLDDQLNSPY----- 68
 QY 218 CTBEGWSCQ--GPFMDSCLSRHSINPYSNRESRIIFSFWNLHIT-----EK 264
 DB 69 -----SCQAEGR-----CLTR-----EVAQLRAENCDLHRRLCSRLCLAEBSR 108
 QY 265 KRTIIPLEAIKRODR-----EVDMEYFGLFTSENKLVIHY 305
 DB 109 QALLESLELPAQDAAGQPPPSQSDKDMKKKKKESEADSEYVHOPFIIFERLKLFEIL 168
 QY 306 CHKRTTKL 314
 DB 169 ---KKDHQL 174

RESULT 19

US-10-097-340-23

Sequence 23, Application US/10097340

Publication No. US20030087250A1

GENERAL INFORMATION:

APPLICANT: John MONAHAN

APPLICANT: Manjula GANNAVARAPU
 APPLICANT: Sebastian HOESCH
 APPLICANT: Shubhangt KAMATKAR
 APPLICANT: Steve G. KOVATS
 APPLICANT: Rachel E. MEYERS
 APPLICANT: Michael MORRISSEY
 APPLICANT: Peter OLANDT
 APPLICANT: Ami SEN
 APPLICANT: Peter VEIBY
 APPLICANT: Gordon B. MILLS
 APPLICANT: Robert C. BAST, Jr.
 APPLICANT: Karen LU
 APPLICANT: Rosemarie SCHMANT
 APPLICANT: Xunel ZHAO
 APPLICANT: Karen GLATT
 TITLE OF INVENTION: Nucleic Acid Molecules and Proteins for The Identification,
 FILE REFERENCE: MRI-030
 CURRENT APPLICATION NUMBER: US/10/097, 340
 CURRENT FILING DATE: 2002-03-14
 PRIOR APPLICATION NUMBER: 60/276, 025
 PRIOR FILING DATE: 2001-03-14
 PRIOR APPLICATION NUMBER: 60/325, 149
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 60/276, 026
 PRIOR FILING DATE: 2001-03-14
 PRIOR APPLICATION NUMBER: 60/324, 967
 PRIOR FILING DATE: 2001/09/26
 PRIOR APPLICATION NUMBER: 60/311, 732
 PRIOR FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: 60/325, 102
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 60/323, 580
 PRIOR FILING DATE: 2001-09-19
 NUMBER OF SEQ ID NOS: 363
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 23
 LENGTH: 633
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-097-340-23

Query Match 4.6%; Score 82.5; DB 9; Length 633;
 Best Local Similarity 20.1%; Pred. No. 36;
 Matches 73; Conservative 48; Mismatches 116; Indels 127; Gaps 18;

QY 37 FQLEGRSRLCLYEDGTETEDYFSPVDNAELVIL-----TLGQAMQGY 82
 DB 65 FSLNEGVRQLKTELGSEFFE-VLQNDLLTKGMVILRDKIRFEGQKLDSLAETWDF 123
 QY 83 SDIRRFSAFHEPQVGLQAQQLCDEQAPORORLLADLIH-----NVS-ONIAET 134
 DB 124 SDVLPMLQALIPYPOGKEPSYRLA-----LLHFNATITLSKLEDALARA 169
 QY 135 RAEDPP-----WFEGLSEFQSGYLRSCSRIRSYREVSYPSTVGAEQAE-- 184
 DB 170 HARVPPIVQMLVLQGVHESRGVTEYLRL--LEFTVQ---KVSPYLGTYGLHSSGPF 224
 QY 185 -----EFLR-----VLGSMQCLRSQVNGS-----YFDGAKG-- 213
 DB 225 THSCITLQQRKAAAMAAALVAGVRRPSPVPAQDPALSTSHPIYDVARHGILQVAGDDR 284
 QY 214 -GSRICTEGWFSC-----QGFPMDSCLSRHSINPYSNRESRIIFSFWNLHIEKK 265
 DB 285 FGRVYV-----FSCCRMPSPSHLDHQRLELYKTLTLOQYVENDTIYFPHGLN---SRN 337
 QY 266 RTIIPLEAIKEDGREVD-----WEYFYFGLFTSENKL 301
 DB 338 KPSLGLWLSAKYKEDRKDGDLTWPRLVYNSKLRSSHLSLPKYWDYR-----KKNLKA 392
 QY 302 VHIY 305
 DB 393 LYVY 396

THIS PAGE BLANK (USPTO)

Tue May 27 09:31:54 2003

us-09-748-451-2.ral

Page 1

Gencore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 15:19:57 ; Search time 21 Seconds

(without alignments)
473.569 Million cell updates/sec

Title: US-09-748-451-2

Perfect score: 1789

Sequence: 1 MLOKPSVKLRALSRPRKFG.....SRVYKQTRLRKQPVKRRQ 338

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

Issued_Patents_AA:*
1: /cgn2-6/ptodata/1/aa/5A_COMB.pep:*
2: /cgn2-6/ptodata/1/aa/5B_COMB.pep:*
3: /cgn2-6/ptodata/1/aa/6A_COMB.pep:*
4: /cgn2-6/ptodata/1/aa/6B_COMB.pep:*
5: /cgn2-6/ptodata/1/aa/6C_COMB.pep:*
6: /cgn2-6/ptodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	1789	100.0	338	US-09-061-702-2	Sequence 2, Appl1
2	89	5.5	466	US-09-610-401-4	Sequence 4, Appl1
3	87	5.4	466	US-09-610-401-3	Sequence 3, Appl1
4	81	5.1	219	US-09-069-023-18	Sequence 18, Appl1
5	89	5.0	331	US-09-061-702-4	Sequence 4, Appl1
6	89	5.0	457	US-08-847-900-3	Sequence 3, Appl1
7	88.5	4.9	3080	US-09-085-1998-9	Sequence 9, Appl1
8	88	4.9	756	US-09-069-023-32	Sequence 32, Appl1
9	87	4.9	1463	US-08-157-005-3	Sequence 3, Appl1
10	86.5	4.8	1463	US-08-747-863-3	Sequence 3, Appl1
11	86.5	4.8	1463	US-09-565-864-3	Sequence 3, Appl1
12	86.5	4.8	219	US-09-069-023-12	Sequence 12, Appl1
13	86	4.8	2296	US-08-286-819A-27	Sequence 27, Appl1
14	85.5	4.8	2296	US-08-980-357-27	Sequence 27, Appl1
15	85.5	4.8	2296	US-08-945-056-2	Sequence 2, Appl1
16	84.5	4.7	373	US-09-069-023-16	Sequence 16, Appl1
17	84	4.7	217	US-09-069-023-31	Sequence 31, Appl1
18	83.5	4.7	239	US-09-413-814-28	Sequence 28, Appl1
19	83.5	4.6	2618	US-09-069-023-20	Sequence 20, Appl1
20	82.5	4.6	299	US-08-659-251-5	Sequence 5, Appl1
21	82.5	4.6	1055	US-08-256-180-5	Sequence 5, Appl1
22	82.5	4.6	1055	PCT-US93-11445-5	Sequence 5, Appl1
23	82.5	4.6	1055	US-08-384-917A-6	Sequence 6, Appl1
24	81.5	4.6	628	US-08-394-912A-6	Sequence 6, Appl1
25	81.5	4.6	628	US-09-333-636-6	Sequence 6, Appl1
26	81.5	4.6	630	US-08-394-177-5	Sequence 5, Appl1
27	81.5	4.6	630	US-08-394-177-5	Sequence 5, Appl1

Applicants only

28	81.5	4.6	630	US-08-394-912A-5	Sequence 5, Appl1
29	81.5	4.6	630	US-09-333-636-5	Sequence 5, Appl1
30	81.5	4.6	870	US-09-172-339-8	Sequence 8, Appl1
31	81	4.5	2818	US-08-510-284-1	Sequence 1, Appl1
32	81	4.5	2818	US-08-411-389-2	Sequence 2, Appl1
33	81	4.5	2818	US-08-449-933-2	Sequence 2, Appl1
34	81	4.5	2818	US-07-966-049A-2	Sequence 2, Appl1
35	81	4.5	2818	US-09-542-331-2	Sequence 2, Appl1
36	81	4.5	2818	US-09-510-791-2	Sequence 2, Appl1
37	80.5	4.5	1001	US-09-060-410-2	Sequence 2, Appl1
38	79.5	4.4	626	US-07-938-782A-2	Sequence 2, Appl1
39	79.5	4.4	626	US-08-630-524-2	Sequence 2, Appl1
40	79.5	4.4	626	PCT-US93-08131-2	Sequence 2, Appl1
41	79	4.4	323	US-09-619-353-6	Sequence 6, Appl1
42	78.5	4.4	384	US-08-707-793A-5	Sequence 5, Appl1
43	78.5	4.4	384	US-08-707-792A-5	Sequence 5, Appl1
44	78.5	4.4	402	US-08-553-703A-1	Sequence 1, Appl1
45	78.5	4.4	402	US-09-006-021-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-061-702-2
Sequence 2, Application US/09061702
Patent No. 6165737
GENERAL INFORMATION:
APPLICANT: Wang, Xiaodong
TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN
TITLE OF INVENTION: APOPTOSIS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/061,702
FILING DATE: Concurrently Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McMillian, Nabeela R.
REGISTRATION NUMBER: P-43,363
REFERENCE/DOCKET NUMBER: UTSD:546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512)418-3000
TELEFAX: (512)474-7577
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 338 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-09-061-702-2
Query Match
Best Local Similarity 100.0%; Score 1789; DB 4; Length 338;
Matches 338; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB
1 MLOKPSVKLRALSRPRKFGVAGRCQEVLRKGLRFPOLPERGSRICLYEDGTETEDYF 60
1 MLOKPSVKLRALSRPRKFGVAGRCQEVLRKGLRFPOLPERGSRICLYEDGTETEDYF 60
QY
61 PSVPDAEALVLTLLGQAMGVYSDIRRLSFAHFPOVGLIOAQQLLCDEQAPORRLA 120

DB 61 PSVPDNEVLTLTGQAMQGVSDIRRLFSFHEPQGLTQAQQLCDDEAPQORLLA 120
QY 121 DLHNVSONIAETRAEDPPMEFEGLESFQSGSYLRSCESTRSYLREVSYPSTYGA 180
DB 121 DLHNVSONIAETRAEDPPMEFEGLESFQSGSYLRSCESTRSYLREVSYPSTYGA 180
QY 181 EAGEEFLRVLSGMCQRLRSQYNGSYFDRGAKGSGRLCTPEGWFSGCGPEMDSCLSRHS 240
DB 181 EAGEEFLRVLSGMCQRLRSQYNGSYFDRGAKGSGRLCTPEGWFSGCGPEMDSCLSRHS 240
QY 241 INPVSNESTRLESTWMLDHIETKKRTIPLVFAIKEDQGRFVDMETFGLLFTSENK 300
DB 241 INPVSNESTRLESTWMLDHIETKKRTIPLVFAIKEDQGRFVDMETFGLLFTSENK 300
QY 301 LVHIVCHKTKTKLNCDSRIRYKQPTRLKRPYRRKQ 338
DB 301 LVHIVCHKTKTKLNCDSRIRYKQPTRLKRPYRRKQ 338
QY 2
-610-401-4
Sequence 4, Application US/09610401
Patent No. 6417336
GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. 6417336uhitro,
APPLICANT: NAKANISHI, Keiko,
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522.1004
CURRENT APPLICATION NUMBER: US/09/610.401
CURRENT FILING DATE: 2000-07-05
PRIOR APPLICATION NUMBER: JP 11-193235
PRIOR FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 466
TYPE: PRT
ORGANISM: Mus sp.
US-09-610-401-4
Query Match 5.5%; Score 99; DB 4; Length 466;
Best Local Similarity 20.18; Pred. No. 0.054;
Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;
QY 27 QEVLRKGLRFPOLPERGSRCLYEDTELTEDYFPPVDNAELVLTGQAMQGVSDIR 86
DB 180 EDIMR---LREKIOEE---MLQREAEASTLQSFROVDNASTLARLDLKRVESTLOEPI- 231
87 RFLSAFHEPQVGLTQA-----QQLCDDEAPQORLLADLHNVSONIAETRAEDPPW 141
QY 232 AFLKHLDEETIOELQAOIOEHVOIDVDSKPDLTALRDVROQY-ESVAKNKIQEAEEW 290
DB 142 F-----GLSFRFSQSGSYLRY-----SCESRIR-----S 166
QY 291 YKSFADLSEANNNALRQAKQESNEYRQVQSLTCEVDALKTNESLERQREMEEN 350
DB 167 YLREVSYPSTYGAEAGEEFLRVLSGMCQRLRSQYNGSYFDRGAKGSGRLCTPEGWFS 226
QY 351 FAVEANAYODTIG-RLODEIOMKEEMARHLREYO-----DILLNV 389
DB 227 QGPEMDSCLSRHSINPVSNESTRIL-----LFTWMLDHI-----IEKRTIIP 272
QY 390 KMALDIEAIYRKIL---EGESRISLPLPFSSILNRETMLSLPIVDTHSKRTL---L 443
DB 273 VEAIKEDQGREVD 285
QY 444 IKVETRDQGVIN 456

Sequence 3, Application US/09610401
Patent No. 6417336
GENERAL INFORMATION:
APPLICANT: MORISHIMA, No. 6417336uhitro,
APPLICANT: NAKANISHI, Keiko,
APPLICANT: SHIBATA, Takehiko
TITLE OF INVENTION: Antibody against cleavage product of vimentin
FILE REFERENCE: 522.1004
CURRENT APPLICATION NUMBER: US/09/610.401
CURRENT FILING DATE: 2000-07-05
PRIOR APPLICATION NUMBER: JP 11-193235
PRIOR FILING DATE: 1999-07-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 3
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-610-401-3
Query Match 5.4%; Score 97; DB 4; Length 466;
Best Local Similarity 20.88; Pred. No. 0.088;
Matches 65; Conservative 46; Mismatches 112; Indels 90; Gaps 15;
QY 27 QEVLRKGLRFPOLPERGSRCLYEDTELTEDYFPPVDNAELVLTGQAMQGVSDIR 86
DB 180 EDIMR---LREKIOEE---MLQREAEASTLQSFROVDNASTLARLDLKRVESTLOEPI- 231
QY 87 RFLSAFHEPQVGLTQA-----QQLCDDEAPQORLLADLHNVSONIAETRAEDPPW 141
DB 232 AFLKHLDEETIOELQAOIOEHVOIDVDSKPDLTALRDVROQY-ESVAKNKIQEAEEW 290
QY 142 F-----GLSFRFSQSGSYLRY-----SCESRIR-----S 166
DB 291 YKSFADLSEANNNALRQAKQESNEYRQVQSLTCEVDALKTNESLERQREMEEN 350
QY 167 YLREVSYPSTYGAEAGEEFLRVLSGMCQRLRSQYNGSYFDRGAKGSGRLCTPEGWFS 226
DB 351 FAVEANAYODTIG-RLODEIOMKEEMARHLREYO-----DILLNV 389
QY 227 QGPEMDSCLSRHSINPVSNESTRIL-----LFTWMLDHI-----IEKRTIIP 272
DB 390 KMALDIEAIYRKIL---EGESRISLPLPFSSILNRETMLSLPIVDTHSKRTL---L 443
QY 273 VEAIKEDQGREVD 285
DB 444 IKVETRDQGVIN 456

RESULT 4
US-09-069-023-18
Sequence 18, Application US/09069023A
Patent No. 6348573
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohito
APPLICANT: Koseki, Takeyoshi
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
FILE REFERENCE: UM-03333
CURRENT APPLICATION NUMBER: US/09/069.023A
CURRENT FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 18
LENGTH: 219
TYPE: PRT
ORGANISM: Mus musculus
US-09-069-023-18
Query Match 5.1%; Score 91; DB 4; Length 219;
Best Local Similarity 25.88; Pred. No. 0.12;
Matches 46; Conservative 21; Mismatches 71; Indels 40; Gaps 7;

QY 14 RSPRRGVAGRSCQEVLRKKGRLRQLEPERGSRILCIYEDTGL-TEDYPSVADNAELVL 72
Db 45 RYVRK-GLTASLDELDK-VLETLRLRGLTYLEEDGTAVDSDFOLLEDTCTMLV 102
QY 73 TLGQAM---QGVVS-----DIRRFLSAFHEPO-----VGLIOAOQL 107
Db 103 EOGGWSRSGMLSTGCLREKPKHSDIARITFDYVKQNPRLFGLNKKAFYGLYSMS 162
QY 108 CDEQAPORRLADLHNVSQNIATETRAEDPWFEGLESFQSGSYLRSCSRIR 165
Db 163 CPOGVRPRVRLRELTRETSQ-----QGLHMLLGISTLRNVVGADR 208

RESULT 5

US-09-061-702-4
Sequence 4, Application US/09061702
Patent No. 6165737

GENERAL INFORMATION:

APPLICANT: Wang, Xiaodong
TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN
TITLE OF INVENTION: APOPTOSIS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/09/061.702
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McMillian, Nabesela R
REGISTRATION NUMBER: P-43363
REFERENCE/DOCKET NUMBER: UTSD:546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512)418-5000
TELEFAX: (512)474-7577
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 331 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-061-702-4

Query Match 5.0%; Score 89; DB 4; Length 331;

Best Local Similarity 22.2%; Pred. No. 0.38;

Matches 78; Conservative 49; Mismatches 146; Indels 78; Gaps 15;

QY 5 PSKVKLRAR-----SPRRGVAGRSCQEVLR-KGC--LRQLPERGSRILCIYEDGT 53
Db 10 PEGEIRTLKPCLLRRNSRQHVAA-SLEDLRKKADILAIKSLRPTVLAIEGT 68
QY 54 EL-TEDYPSVADNAELVLTLGQAMQGYSDIRRLSAFHEPOVGLIOAOQLCDEQA 112
Db 69 IYDDDDYFLCLPSNRKFKALASNKW-----AYNSDGTAMISDSVDDET 116
QY 113 PORQRLADLHNVSQNIATETRAEDPWFEGLESFQSGSYLRSCSRIR 172
Db 117 DSGAGL-----KKKNVARELK-----EDLSSTILLSEEDLMLVDAPCSDLADELR 162
QY 173 STPSYVGADEAEFLRVLGSMQRLNSMOYNGSYFDRGAKGSGRLCTPEGWFSGCGPFD 231

Db 163 QSCATV-QRLQHTLQVLDQREEVKOSKOLLQLYLALEKESLSKOE---ESKAAGE 218
QY 232 ---MDCSLRSRINPYNSRESRLFTSWNDHIEKRTIPTVLAIEKEDGREYDWE 287
Db 219 EYDAVDTGMSRET-----SSDVALAS---HIL-----TALREKQAPE----- 252
QY 288 YFYGLFTSENKLVHIVCHKTKTKINCDSRIRYKQPTLRKKQPVRRQ 338
Db 253 ---LSISSQDELVTYKEDPKALAVANMIDIKKTEVYQENCERELARLQO 299

RESULT 6

US-08-847-900-3
Sequence 3, Application US/08847900
Patent No. 5963261

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL NEURONAL EXTRACELLULAR MATRIX
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: Fastspo for Windows Version 2.0
CURRENT APPLICATION DATA:
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PR-0282 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 457 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 442368
US-08-847-900-3

Query Match 5.0%; Score 89; DB 2; Length 457;

Best Local Similarity 21.8%; Pred. No. 0.67;

Matches 60; Conservative 40; Mismatches 71; Indels 104; Gaps 16;

QY 44 SRLCIYEDGTETEDYPSVADNAELVLTLGQAMQGYSDIRRLSAFHEQ---VGLI 100
Db 9 SLAVLVMGTEILO---VLPTNPE-----ESMOYV-----SSAQDSRGRCICTV 49
QY 101 QAAQQLCDEQAQRO-RLIADLHNVSQNIATETRAEDPWFEGLESF 149
Db 50 VAPQWQTCRSDATKQLRLEKYVQNMOSIEVLDRTORDIQYKEMENOMKGLSEKFR 109
QY 150 -----OSKGYLR-----YSCSR-IRSYLREVSYPST----- 177

Db 110 QVERSHKOHAKRKAKEKAKADELRLPLVLEEKAKAKVLOREKREYONITSLVNELOE 169
 QY 178 -VGAEOEELRLVLSGMOQLR-----SMOYNGSYFDGANGSGRL- 217
 Db 170 EIGAYDDELQSRVSNLEERLRACQKLAGCKLNGISDPYVKXSGSF-----GSMNT 223
 QY 218 --CTPEG---WFSCQCFPMDSCLSRISINPYSN 246
 Db 224 DPLAPEGDNRYW-----MDGYHNNRFVREKX 251

RESULT 7
 5223423-4

Patent No. 5223423
 APPLICANT: FRANCHINI, GENOVEFFA, WONG-STAAI, FLOSSIE;
 GALLO, ROBERT

TITLE OF INVENTION: CHARACTERIZATION OF REPLICATION COMPETENT
 HUMAN IMMUNODEFICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/15Y

NUMBER OF SEQUENCES: 4

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/331.212
 FILING DATE: 03-31-1989

ID NO: 4

LENGTH: 3080

Query Match 4.9%; Score 88.5; DB 6; Length 3080;
 Best Local Similarity 27.8%; Pred. No. 13;
 Matches 40; Conservative 20; Mismatches 55; Indels 29; Gaps 7;

QY 16 PRKGVAGRSQEQELRKGLRGLPQLPERSRLCTLEDGELTDFPSPDNAE---LYL 71
 Db 861 PHPGLAKKRRTYLDVGDVAFSIP-----LIEDRQYAFPLPSV-NNAEFGKRYIY 912
 QY 72 LTLGQANQ---YVSDIRFLSAFHE--PQVGLQAQAQQLCDEQAPQROQL----- 118
 Db 913 KVLPGQMGKSPAIQYIMRQVLEPFRKANPDVITIVQMDLLI---ASDRIDLEHDKVYL 969
 QY 119 -LADLLHNVSONIAETRAEDPPW 141
 Db 970 QLKELNLGLGFSTDEKFOKDPY 993

RESULT 8

US-09-085-199B-9
 Sequence 9, Application US/09085199B
 Patent No. 6235879

GENERAL INFORMATION:

APPLICANT: Hayden, Michael R.

APPLICANT: Hacken, Abigail

APPLICANT: Chopra, Vikramjit Singh

APPLICANT: Kalchman, Michael

TITLE OF INVENTION: Apoptosis Modulators That Interact with the

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Oppedahl & Larson

STREET: PO Box 5270

CITY: Frisco

STATE: CO

COUNTRY: USA

ZIP: 80443-5270

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 Inch, 1.44 Kb storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS DOS 5.0

SOFTWARE: WordPerfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/085,199B

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Larson, Marina T.
 REGISTRATION NUMBER: 32038
 REFERENCE/DOCKET NUMBER: UBC-P-0130S2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (970) 668-2050
 TELEFAX: (970) 668-2052
 INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 756

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: no

ORIGINAL SOURCE:

ORGANISM: mouse

FEATURE:

OTHER INFORMATION: Huntington-interacting protein

Query Match 4.9%; Score 88; DB 4; Length 756;
 Best Local Similarity 22.4%; Pred. No. 1.7;
 Matches 72; Conservative 47; Mismatches 104; Indels 98; Gaps 15;

QY 22 AGRSQEVLRKGLRGLPQLPERSRLCTLEDGELTDFPSPDNAELVLTGQANQGY 81
 Db 169 AQRSLTELRKAQANEQ---RYSK--LREKYSSELVONHADLLRNAB----- 210
 QY 82 VSDIRFLSAFHEPQVGLQAQQL-----CDEQAPQROBLADLLHNVSONIAETRA 136
 Db 211 ---VTKQVSARQAQVLDLERKELADSPAVSQQAQKQEOQDVLNKLHELAT--SKQ 266
 QY 137 EDPWFEGLESRFQSKGYLYRSCESRIRSYLRVSSYPSTVGAEOEF----- 186
 Db 267 ELQVLSNLETSASQEKWILQIAE-----LEKQCSLATAVQAQREELSLALDQEST 320
 QY 187 -LRYLG---SMCOILRSQNGVSYFDGA-----KGSRLCTPEGFSQGPFD----- 231
 Db 321 QIKLAGAESMCQGVQKQKRTLLAGIRKAAERIEQIALSQLEETP-LISCAGSTDHLISK 379
 QY 232 ---MDCLSR-----HSINPYSNRSRLIFSTWNLHIIRKRTIY 269
 Db 380 VSSVSSCLEQLEKNGSQYLACPEDISLHLSIT-----LALHLTGDIQGSATSL 430
 QY 270 -----PTVEAIKEDDRE 283
 Db 431 RAPEPADSLTEACR-QYGRE 450

RESULT 9

US-09-069-023-32
 Sequence 32, Application US/09069023A
 Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohiro

APPLICANT: Koseki, Takeyoshi

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

FILE REFERENCE: UM-03333

CURRENT APPLICATION NUMBER: US/09/069,023A

CURRENT FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 38

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 32

LENGTH: 331

TYPE: PRT

ORGANISM: Mus musculus

US-09-069-023-32

Query Match 4.9%; Score 87; DB 4; Length 331;
 Best Local Similarity 22.4%; Pred. No. 0.62;
 Matches 81; Conservative 50; Mismatches 131; Indels 100; Gaps 18;

```

QY 5 PMSVKLRALR-----SPRKEGVAGRSQCEVLR-KGC--LRFOLPGRSRLCTYEDGT 53
DB 10 PSCGETRLKCOLLRNNYSRROHGVAA-SCLEDLRKADIIAIKSLTPVTLVLAEDGT 68
QY 54 EL-TEDYPSVDNAEVLVLLTLCQAMOGYSDIRRLRSLFHPQGLLQAOQL-----C 108
DB 69 IYDDDDYFLCPSNTKFKVALASNEKW-----AYNSDGTAMISQESFVDYET 116
QY 109 DQAPORRLADLLHNVSQIAETRAE-----DPPWEGESRFSKSGYLRSGES- 162
DB 117 DSGAGLKKKNVARKQLEKDLISLILSLEDLQMLVDRPCDLAQE-----LNGSCATV 168
QY 163 -RHSYLRVSSYSTYGAENQOEFLRVLSMGQRKRSQYSGSFEDGAGKSGRLCTPE 221
DB 169 ORLQHTLQOVL-----QREVR-----QSQQLQLVLALEKESLSKOE 210
QY 222 GMSGCGPPD-----MDSLSRHSINFSNRESRLIFSTWINDITIEKKRTIPLVEAI 276
DB 211 ---ESKAAGEEVDADVDTGISRET-----SSDVALLS---HTL-----TAL 245
QY 277 KQDGRVDMVEFYGLFTSBNLKVHIVCHKTKHLKNCDSRLTKPKQRLKRPVK 336
DB 246 RERQAPE-----LSLSSODELVTKEDPKALAVANMDIKTETVQEAECBELALRL 297
QY 337 RQ 338
DB 298 QQ 299

RESULT 10
US-08-157-005-3
; Sequence 3, Application US/08157005
; Patent No. 5620691
; GENERAL INFORMATION:
; APPLICANT: Mensvoort, Gert
; APPLICANT: Terpstra, Catharinus
; APPLICANT: Pol, Johannes M
; APPLICANT: Moorman, Robertus J
; TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,005
; FILING DATE: 26-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91201398.4
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200781.0
; FILING DATE: 18-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL92/00096
; FILING DATE: 05-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Morai, Thomas F
; REGISTRATION NUMBER: 16,579
; REFERENCE/DOCKET NUMBER: 44819
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550

```

```

; TELEFAX: (212) 315-1931
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1463 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-157-005-3

Query Match 4.8%; Score 86.5; DB 1; Length 1463;
Best Local Similarity 22.1%; Pred. No. 6.6;
Matches 63; Conservative 41; Mismatches 96; Indels 85; Gaps 15;

QY 100 IQAAQQLCDEQAPORRLADLLHNVSQIAETRA-----EDPPWEGEL 145
DB 558 IEAGRLV-----PNNDRIILAALVHMKQANASRYASAAILMDSACIDHDEWYEDL 612
QY 146 ---ESRFSKSGY-----LRYSCSRISYLRVSSYSTYGAENQOEFLRVLG-SWCQ 195
DB 613 ICGIARCARQDDYSPFGPAFFWSMWEKLRSH-NEGKKFRGICIDAKADYASAGIDL- 670
QY 196 RLRSQYNGSYFDR-----GAKGGSRLCTPEGMSGCGPF-----DMSCLSRHSIN 242
DB 671 ---LFSSHFHQCHXYTLSCGHHAGSKES-----QCSPPVAGACRSPDLAVLKQIPYK 719
QY 243 PYSNRESRLIFSTWIND-----HIEKKRTIPLVEAIKEDGREVDMEVEYGLLF 294
DB 720 PPRVYLMKYGNKTTALDQGRYQSRGLVAVKRG-----AGNEVDLSQXDQ-VVPLLP 772
QY 295 TSENKLVHIVCH-----KRTTHKLN--CDPSRLTKP 324
DB 773 TCDIMWVAVCNVLSKFIYQPSGSKITWLSQVDDVVIYXP 817

RESULT 11
US-08-747-863-3
; Sequence 3, Application US/08747863
; Patent No. 6197310
; GENERAL INFORMATION:
; APPLICANT: Mensvoort, Gert
; APPLICANT: Terpstra, Catharinus
; APPLICANT: Pol, Johannes M
; APPLICANT: Moorman, Robertus J
; TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Traak, Britt & Rosaa
; STREET: 525 South 300 East
; CITY: Salt Lake City
; STATE: Utah
; COUNTRY: USA
; ZIP: 84111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: WINDOWS95
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/747,863
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/157,005
; FILING DATE: 26-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91201398.4
; FILING DATE: 06-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 92200781.0
; FILING DATE: 18-MAR-1992

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/NL92/00096
FILING DATE: 05-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Turner, Allen C.
REGISTRATION NUMBER: 33041
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (801) 532-1922
TELEFAX: (801) 531-9168
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1463 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-747-863-3

Query Match 4.8%; Score 86.5; DB 4; Length 1463;
Best Local Similarity 22.1%; Pred. No. 6.8;
Matches 63; Conservative 41; Mismatches 96; Indels 85; Gaps 15;

100 IQAAQQLCDQEAQORQLADLHNVSQNIATRA-----EDPPWFEGL 145
558 IEAGRLV-----PNDRIILALAYHMAQNASEYASAAIIMDSCACIDHDEWYEDL 612
QY 146 ---ESRFQSKSGY-----LRYSCESRIRSYLREVSYPTVGAEAQEFRLVIG-SMCQ 195
DB 613 ICGIARCARQDGYFPGAFPMWMEKLRSH-NEGKFRHCGICDADKADVASAGLDLC- 670
QY 196 RLRSQNGSYFDR-----GAKGSRILCTPGWFSQCGP-----DMDSCLSRHSIN 242
DB 671 -----LFHSHFHCHXVTLSCGHAGSKES-----QCSPPVAGRSPLDAVLKQIPYK 719
QY 243 PYSNRESILFSTNLD-----HIEKRTIIPFLVEAIKEDQREVDMEYFGLLF 294
DB 720 PPRIVIMKVGKKTALDGRYQSRRLVAVKRG-----AGNEVDLSDXDYQ-VPLLP 772
QY 295 TSENILKVHYCH-----KKTTHKLN-CDPSRIYRP 324
DB 773 TCKDINMKVACNVLLSKFIYGPSPGSKTWTLLSQVDDDVITYP 817

RESULT 12

US-09-565-864-3
Sequence 3, Application US/09565864
Patent No. 6455245
GENERAL INFORMATION:

APPLICANT: Wensvoort, Gert
Terpstra, Catharinus
Pol, Johannes M
Moortman, Robertus J
Meulenbergh, Johanna J

TITLE OF INVENTION: CAUSITIVE AGENT OF THE MYSTERY SWINE DISEASE,
VACCINE COMPOSITIONS AND DIAGNOSTIC KITS

NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.24
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/565,864
FILING DATE: 05-MAY-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/157,005
FILING DATE: <Unknown>
APPLICATION NUMBER: EP 91201398.4
FILING DATE: 06-JUN-1991
APPLICATION NUMBER: EP 92200781.0
FILING DATE: 18-MAR-1992
APPLICATION NUMBER: PCT/NL92/00096
FILING DATE: 05-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Moran, Thomas F
REGISTRATION NUMBER: 16,579
REFERENCE/DOCKET NUMBER: 44819
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 315-1931
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1463 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-565-864-3

Query Match 4.8%; Score 86.5; DB 4; Length 1463;
Best Local Similarity 22.1%; Pred. No. 6.8;
Matches 63; Conservative 41; Mismatches 96; Indels 85; Gaps 15;

100 IQAAQQLCDQEAQORQLADLHNVSQNIATRA-----EDPPWFEGL 145
558 IEAGRLV-----PNDRIILALAYHMAQNASEYASAAIIMDSCACIDHDEWYEDL 612
QY 146 ---ESRFQSKSGY-----LRYSCESRIRSYLREVSYPTVGAEAQEFRLVIG-SMCQ 195
DB 613 ICGIARCARQDGYFPGAFPMWMEKLRSH-NEGKFRHCGICDADKADVASAGLDLC- 670
QY 196 RLRSQNGSYFDR-----GAKGSRILCTPGWFSQCGP-----DMDSCLSRHSIN 242
DB 671 -----LFHSHFHCHXVTLSCGHAGSKES-----QCSPPVAGRSPLDAVLKQIPYK 719
QY 243 PYSNRESILFSTNLD-----HIEKRTIIPFLVEAIKEDQREVDMEYFGLLF 294
DB 720 PPRIVIMKVGKKTALDGRYQSRRLVAVKRG-----AGNEVDLSDXDYQ-VPLLP 772
QY 295 TSENILKVHYCH-----KKTTHKLN-CDPSRIYRP 324
DB 773 TCKDINMKVACNVLLSKFIYGPSPGSKTWTLLSQVDDDVITYP 817

RESULT 13

US-09-069-023-12
Sequence 12, Application US/09069023A
Patent No. 6348573
GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohito

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS

FILE REFERENCE: UM-03333

CURRENT APPLICATION NUMBER: US/09/069,023A

CURRENT FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 38

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 12

LENGTH: 219

TYPE: PRT

ORGANISM: Homo sapiens

Query Match 4.8%; Score 86; DB 4; Length 219;
Best Local Similarity 36.4%; Pred. No. 0.42;

Matches 24; Conservative 11; Mismatches 29; Indels 2; Gaps 2;
OY 14 RSPKRGVAGRCQCVLAKGCLRPQLPERSRLCYEDTEL-TEDYPSYDPAALYL 72
DB 44 RSSRR-GVASSLDELISKILDAIVATGLVLYLEDDQTVDTTEFFQITLGDTHRMIL 102
OY 73 TLGQAW 78
DB 103 ERQOKW 108

RESULT 14
US-08-286-819A-27
Sequence 27, Application US/08286819A
Patent No. 5871910

GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-286-819A-27

Query Match 4.8%; Score 85.5; DB 2; Length 2296;
Best Local Similarity 26.2%; Pred. No. 17;
Matches 33; Conservative 17; Mismatches 41; Indels 35; Gaps 6;
OY 156 LRYSCSS-RIRSYLAEVSSYSTVGAAQGEFFRVAGSMCQRLRSMQV---NGSY-----206
DB 1314 LRPMSIRIRITLISATKY-----RSKHEFLSLIGAV--HLIPMSWLKHKQGNAAVPH 1366
OY 207 --FDGAKGSRICTPCEGFMSCQGFPMDCSLRHSINYSNRRESRIILFTWNLDTIEK 264
DB 1367 WMYKRRKSEFTLINPKQNLINLYLNFKKELTSHRIRP-----11PS 1409
OY 265 KRTIIP 270
DB 1410 KRCVIP 1415

RESULT 15
US-08-980-357-27
Sequence 27, Application US/08980357
Patent No. 6013508

GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPROTEINS, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 2296 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein.
US-08-980-357-27

Query Match 4.8%; Score 85.5; DB 3; Length 2296;

Best Local Similarity 26.2%; Pred. No. 17;

Matches 33; Conservative 17; Mismatches 41; Indels 35; Gaps 6;

156 LRYSES-RIRSYLREVSSYSTGAEAEFLRYLGSRCORLSMOY-----NSGY----- 206

1314 LRFRCRSRIRITISATRYK-----RESKEHFLSLGAV--HLIPMSLKRKTGMWVPH 1366

207 --FDGANGSRLCTPEGWFCQGPEDMDSCLSRHSINPYSNRESRIILFTWNLDHIEK 264

1367 WMYKERRSFSTLAPKKNLIINFLNFKECTLSHRIRP-----ILPS 1409

265 KRITIP 270

1410 KRCVIP 1415

RESULT 16

US-08-945-056-2

Sequence 2, Application US/08945056

Patent No. 6077994

GENERAL INFORMATION:

APPLICANT: Coupland, George M.

APPLICANT: Puttill, Joanna J.

TITLE OF INVENTION: Genetic control of flowering

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon & Vanderhye PC

STREET: 8th floor, 1100 No. 6077994th Glebe Road

CITY: Arlington

STATE: Virginia

COUNTRY: United States of America

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/945,056

FILING DATE: 20-OCT-1997

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/G895/02561

FILING DATE: 01-NOV-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9422083.7

FILING DATE: 02-NOV-1994

ATTORNEY/AGENT INFORMATION:

NAME: Ms Mary J Wilson

REGISTRATION NUMBER: 32,955

REFERENCE/DOCKET NUMBER: 620-17

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 816-4000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-945-056-2

Query Match 4.7%; Score 84.5; DB 3; Length 373;

Best Local Similarity 22.6%; Pred. No. 1.4;

Matches 47; Conservative 26; Mismatches 64; Indels 71; Gaps 10;

153 SGYLRACSESRIRSYLREVSSYSTGAEAEFLRYLGSRCORLSMOYGSYFDGRK 212

36 SAYLCHSCDAQVHSANRYAS-----RHKRYRCES-CEAPAP----- 71

213 GGSRLCTPEGWFCQGPEDMDSCLSRHSINPYSNRESRI-----LFSTWNLDHIE 263

72 --AFLEADSDASLC-----TACSEVHSANPLARHQRVYILPISGSEFSMTTHQS 123

264 KRITIPPLVEAIKRODREVD-----MEY-----FYGLTSTENIKLV--- 302

124 EKTWDPKRLVVDDEEGEDDKAEVSWLFPNSDKNNNNQNGLLFSDYINLVYXN 183

303 ----HIVCHKRTKHLNDPSRIYKPR 326

184 SMDYKFTGEYSQHONCS-----VPR 206

RESULT 17

US-09-069-023-16

Sequence 16, Application US/09069023A

Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohiro

APPLICANT: Koseki, Takeyoshi

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

FILE REFERENCE: UM-03333

CURRENT APPLICATION NUMBER: US/09/069,023A

CURRENT FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 38

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 16

LENGTH: 217

TYPE: PRT

ORGANISM: Mus musculus

US-09-069-023-16

Query Match 4.7%; Score 84; DB 4; Length 217;

Best Local Similarity 36.4%; Pred. No. 0.68;

Matches 24; Conservative 11; Mismatches 29; Indels 2; Gaps 2;

14 RSPRRGVAGRSCEVLRKGLRQLPERSRLCYEGTEL-TEDYPSVDAEALVLL 72

44 RSSRR-GYMASSLDLISKTIDVLYITGLVTLVEDGIVADTEEPOTLRDTHFWTL 102

73 TLQAM 78

103 EKGQKW 108

RESULT 18

US-09-069-023-31

Sequence 31, Application US/09069023A

Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohiro

APPLICANT: Koseki, Takeyoshi

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

FILE REFERENCE: UM-03333

CURRENT APPLICATION NUMBER: US/09/069,023A

CURRENT FILING DATE: 1998-04-27

NUMBER OF SEQ ID NOS: 38

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 31

LENGTH: 239

TYPE: PRT

ORGANISM: Mus musculus

US-09-069-023-31

Query Match 4.7%; Score 83.5; DB 4; Length 239;

Best Local Similarity 28.8%; Pred. No. 0.89;

Matches 23; Conservative 19; Mismatches 35; Indels 3; Gaps 2;

QY 3 ORPKSVKLRALSPRKT--GVAGRSQEVLRKGLRFPOLPERGSRCLYEDGTET--TEDY 59
DB 38 ETPRRAPCRVSTARKRKGMASLEIDLNVODILKOKPPSLVLEBDGTLVETEE 97
QY 60 PPSVDPNAELVLTITLGOAW 79
DB 98 FOALAKDTMFVLLKQKWK 117

RESULT 19

US-09-413-814-28
Sequence 28, Application US/09413814

Patent No. 6225064

GENERAL INFORMATION:

APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH

APPLICANT: Bristol-Myers Squibb, CO.

APPLICANT: Bayer, Stefan

APPLICANT: Bloecher, Helmut

APPLICANT: Brandt, Petra

APPLICANT: Cino, Paul M

APPLICANT: Dougherty, Brian A

APPLICANT: Goldberg, Steven L

APPLICANT: Hottle, Gerhard

APPLICANT: Mueller, Joachim

APPLICANT: Reichenbach, Hans

TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polypeptide or

FILE REFERENCE: PCT/US 99/23535

CURRENT APPLICATION NUMBER: US/09/413,814

EARLIER FILING DATE: 1999-10-07

EARLIER APPLICATION NUMBER: DE 198 46 493.2

NUMBER OF SEQ ID NOS: 107

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 28

LENGTH: 2618

TYPE: PRT

ORGANISM: Sorangium cellulosum

US-09-413-814-28

Query Match 4.7%; Score 83.5; DB 4; Length 2618;

Best Local Similarity 23.0%; Pred. No. 35;

Matches 54; Conservative 25; Mismatches 89; Indels 67; Gaps 7;

QY 40 PERGSRCLYEDGTETEDYPSVDPNAE-----LYLTLTGOAWOGYVDIRFLSA 91

DB 2062 PEROSPRLARHETELGRHLLALRQLAEVGAPEVRLAVTTGLAIGSEELR----- 2116

92 FHEPQVGLQAAQOLCDEQAFORRLIADILHNVSQNAETRAEDPPWEGLE----- 146

2117 ---PEVGLRGVPRVLPFEFENLRLLT-----DIDSADPTWRSGCCPLIRE 2160

QY 147 -----SRQSKSGYLRYSCESRIRSYLREVSYPSTVGAEOEFLRVIG 191

DB 2161 MGAAPGEETALRGTSRWELGEPVGGTSTISSRLREGVYLITG-----LGLIG 2213

QY 192 SMCQRLSMQYNGSYFDRNGAKGSRICTPEGWFSQCGP-----DMDSCLSR 238

DB 2214 LALARHLARKYRATLITAGRGP---ARELHQAPAEFVPAALQOMECCAR 2265

RESULT 20

US-09-069-023-20
Sequence 20, Application US/09069023A

Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohiro

APPLICANT: Koseki, Takayoshi

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS

FILE REFERENCE: DM-03333

CURRENT APPLICATION NUMBER: US/09/069,023A

CURRENT FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 299
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-069-023-20

Query Match 4.6%; Score 82.5; DB 4; Length 299;

Best Local Similarity 29.7%; Pred. No. 1.6;

Matches 22; Conservative 16; Mismatches 35; Indels 1; Gaps 1;

QY 6 KSVKLRALSPRKFEVAGRSQEVLRKGLRFPOLPERGSRCLYEDGTET--TEDYPSVP 64

DB 12 KPEKYVDVYRNIRIKKAVCASISLEIRSVAKFEKCDHPTILDSGTEIDDEYFRTLD 71

QY 65 DNAELVLTITLGOAW 78

DB 72 ENTELVAVPEGEHW 85

Search completed: May 26, 2003, 15:24:23
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 15:14:07 ; Search time 37 Seconds
(without alignments)
878.201 Million cell updates/sec

Title: US-09-748-451-2

Perfect score: 1789
Sequence: 1 MGRKSVKRLALRSPKRF.....SRVYKQRLRKQPVRRRQ 338

Indexing table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 283224

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	108.5	6.1	1920	2	A53188
2	105.5	5.9	571	2	pericentrin - mouse
3	104	5.8	1778	2	NBS-LRR type resist
4	99.5	5.6	2629	2	apollipoprotein B -
5	99	5.5	466	2	apollipoprotein B -
6	98.5	5.5	466	2	vimentin - human
7	98.5	5.5	684	2	vimentin - mouse
8	98	5.5	448	2	helicase IV (75 kd
9	98	5.5	448	2	vimentin - Chinese
10	98	5.5	464	1	vimentin - golden
11	98	5.5	594	2	slp(w7) - mouse (f
12	96.5	5.4	571	2	acetylactate synth
13	93.5	5.2	397	2	COI intron 10 prot
14	92.5	5.2	1054	1	HIV-1 retropepsin
15	92	5.1	1735	2	sex-limited protei
16	92	5.1	1736	2	sex-limited protei
17	91.5	5.1	925	2	oxido/reductase ir
18	91.5	5.1	925	2	ct234, hypothetical
19	91.5	5.1	925	2	conserved hypoteth
20	91.5	5.1	925	2	CT234, hypothetical
21	91	5.1	460	2	vimentin - chicken
22	91	5.1	776	2	hypothetical prote
23	90	5.0	1044	2	hypothetical prote
24	90	5.0	1211	2	probable calphosin
25	89.5	5.0	763	2	glucose-6-phosphat
26	89.5	5.0	4116	2	calo protein - fru
27	89	5.0	457	2	neuronal olfactome
28	89	5.0	609	2	hypothetical prote
29	88.5	4.9	458	2	desmin - African c

30	88.5	4.9	1032	2	S12153
31	88.5	4.9	1306	2	T13592
32	88.5	4.9	1738	1	A24558
33	87.5	4.9	1035	1	GNLJGC
34	87.5	4.9	1242	2	S78061
35	87	4.9	1066	2	T45283
36	87	4.9	1066	2	T41099
37	86.5	4.8	1121	2	T38127
38	86	4.8	437	2	C75632
39	86	4.8	458	2	A43549
40	86	4.8	1027	2	S37711
41	86	4.8	1027	2	B43549
42	86	4.8	1571	2	T14155
43	86	4.8	2895	2	H83362
44	85.5	4.8	987	2	T51360
45	85.5	4.8	1744	1	C4HU

ALIGNMENTS

RESULT 1

A53188
pericentrin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: A53188
R:Boxsey, S.J.; Stein, P.; Evans, L.; Calarco, P.D.; Kirschner, M.
Cell 76, 639-650, 1994
A:Title: Pericentrin, a highly conserved centrosome protein involved in microtubule o
A:Reference number: A53188; MIMD:94170365; PMID:8124707
A:Accession: A53188
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1920 <BOX>
A:Cross-references: GB:005823; NID:9458667; PID:MAL1786.1; PID:9458668
C:Keywords: coiled coil

Query Match

Best Local Similarity 23.9%; Pred. No. 1.7;
Matches 56; Conservative 45; Mismatches 90; Indels 43; Gaps 11;

QY	2	LQKPKSVKRLALRSPKRFVAGRSQEVY----	RGK-----LRPOLPGRSGRICTY--	ED	51
DB	142	LQKPKSVKRLALRSPKRFVAGRSQEVY----	RGK-----LRPOLPGRSGRICTY--	ED	51
QY	52	GTELTEDYFPSPVDNAELVLTQAOAG-----	YVSDIRRLSAFEPQVGLIAAOQL	107	
DB	198	AAETKRLRSEMEKNAQ--TETLKQDWESERELCTE	MLROELSLKQSEMEGLQSQFOKE	256	
QY	108	CDEQAPQORRL-----ADLIHVNSONIAA-----	ETRADPPMFEGLSRFSQSG	154	
DB	257	LSEQKVELEKIFQAKHEAVSLKNEAQHQAIRKIQ	EDIQSFHCQYLDLQKPRERK	316	
QY	155	YLRSCSRIRSYLRVSSYSTVGAEQEFELVLSMCQ	RLRS--MQVNGSY	206	
DB	317	AKELLETTLQASY-----EDLKAQGE--	ILLMSQLESKNTRRELNSW	361	

RESULT 2

T02213
NBS-LRR type resistance protein - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 17-Nov-2000
C:Accession: T02213
R:Heister, D.; Kutt, J.; Laurie, D.A.; Yano, M.; Sasaki, T.; Devos, K.; Graner, A.;
Proc. Natl. Acad. Sci. U.S.A. 95, 370-375, 1998
A:Title: Rapid reorganization of resistance gene homologues in cereal genomes.
A:Reference number: Z14623; MIMD:98081880; PMID:9419382
A:Accession: T02213
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-571 <LEI>

A:Cross-references: EMBL:AF032688; NID:g2792219; PIDN:AA06985.1; PID:g2792220
 C:Genetics:
 A:Gene: r1
 C:Superfamily: disease resistance protein RPS2; leucine-rich alpha-2-glycoprotein repeat

Query Match 5.9%; Score 105.5; DB 2; Length 571;
 Best Local Similarity 20.7%; Pred. No. 0.62;
 Matches 69; Conservative 38; Mismatches 97; Indels 129; Gaps 13;

QY 43 GSRLCYEDGTELEDPYPSV-----PDNAELVLTLLGAMOGYSDIR--FLSA 91
 DB 240 GSLLC-----TKDEPDMKNVLRSEIWEIPEKNNILPALRLSYNHLPAILLKRCAPFSV 294
 QY 92 FHEPOV-----GLIQAQQLCDE-----QAPORQLADLL 123
 DB 295 FHKYVPEKETLVQIWMALGFIQSPGRTEIELSGSYDELGRSFFHHGGGYMHAM 354
 QY 124 HNVSQNIAM--TRAEDPWFEGLESRFQSGYRSCSRIHSRYLAEVSYPSTVAE 181
 DB 355 HDLAQSVMDCLRLDPP-----NSSSTRSSRHLSFSCNHSRT----- 395
 QY 182 AQEELRYLQSGMQRRLRSMQYNGSYFDGANGSRLCTPEGWFCGPFDMDSLSRHSI 241
 DB 396 SEDEFL-----GPKKARTLLLLNGYKSRTPSPD----- 425
 QY 242 NPYSNRESRIIFSTWNLDHIEKKRTIIPVLEAIKEDGREDVMEF----- 289
 DB 426 -----LFELMLRYHLVLELNRRDITELPDSI-----GNLKMRLYLNLSGTGITVPS 471
 QY 290 -YGLFETSENKLVIHYCHKKTTHKLNCDPSRI 321
 DB 472 SIGRLFNQITLKL-----KNCVLEICIPSI 497

RESULT 3

apolipoprotein B - pig (fragments)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 13-Aug-1999
 C:Accession: J030382; I46567
 R:Maeda, N.; Ebert, D.L.; Doers, T.M.; Newman, M.; Hasler-Rapacz, J.; Attie, A.D.; Rapacz, Gene 70, 213-229, 1988
 A:Title: Molecular genetics of the apolipoprotein B gene in pigs in relation to atherosclerosis
 A:Reference number: J030382; MUID:89100006; PMID:2905687
 A:Accession: J030382

A:Molecule type: DNA
 A:Residues: 1-1778 <MAE>
 R:Maeda, N.; Ebert, D.L.; Doers, T.M.; Newman, M.; Hasler-Rapacz, J.O.; Attie, A.D.; Rapacz, Gene 69, 213-229, 1988
 A:Title: Molecular genetics of the apolipoprotein B gene in pigs in relation to atherosclerosis
 A:Reference number: I46567
 C:Accession: I46567
 C:Superfamily: translated from GB/EMBL/DBDUB

Query Match 5.8%; Score 104; DB 2; Length 1778;
 Best Local Similarity 19.5%; Pred. No. 3.8;
 Matches 66; Conservative 52; Mismatches 123; Indels 98; Gaps 14;

A:Cross-references: GB:M22646; NID:g164366; PIDN:AAA30996.1; PID:g164366
 A:Accession: I46568
 A:Status: preliminary; translated from GB/EMBL/DBDUB
 A:Molecule type: DNA
 A:Residues: 239-1778 <MA3>
 A:Cross-references: GB:M22647; NID:g164367; PIDN:AAA30997.1; PID:g164370
 C:Comment: Apolipoprotein B is the predominant protein component of the low-density lipoprotein
 C:Genetics:
 A:Gene: apob
 A:Introns: 39/3; 88/3; 159/2; 238/3; 1133/1; 1171/2; 1232/3
 A:Note: the list of introns may be incomplete
 C:Superfamily: apolipoprotein B
 C:Keywords: atherosclerosis; cholesterol metabolism; LDL; lipid binding; lipoprotein

Query Match 5.8%; Score 104; DB 2; Length 1778;
 Best Local Similarity 19.5%; Pred. No. 3.8;
 Matches 66; Conservative 52; Mismatches 123; Indels 98; Gaps 14;

QY 36 RFOLEP-----GSRIC-----LY-----EDGTELEDYFPSPVNAEL----- 69
 DB 1404 RFOLEPGRANNTGDELCKMMVTEVGEVLQIYSKIHSGLEILSYFODLMERSKLNALK 1463
 QY 70 -----VLTLGAMOGYSDIRRFSAFHEPOVGLIQAQQLC----- 108
 DB 1464 IKTFPDSKYQVLTVDVSEYGRQLKSLSDVQALSDLSINITTELSLQIFLECFIOEI 1523
 QY 109 -DEQAPORQLADLLHNVSQNIAMETRAEDPWFEGLESRFQSGYRSCSRI-- 164
 DB 1524 EEEFLRLKEKRLADPINDIQNINNTFTVAPLGRRLKENLDSFGMLNFIQNTLMEA 1583
 QY 165 RSYLREVSST-----PSYVG-----AEQEEFLRYLQSGMQRRLRSMQYNGSYFDR 209
 DB 1584 SOELQOLHQYIKALRKEFYDPSMGMVTKYYLEKEKVINLTKNLVDYK--DFHSKYTVS 1641
 QY 210 GAKGSRCTPEGWFCGPFDMDSCLS-----RHSINPYSNRESRIIFSTWNL 260
 DB 1642 ATDFASQLSQVEQVQVQ--DIQRYLSITLADADGKGKELAEISSRAQETI--KSAV-- 1695
 QY 261 IIEKKRTIIPVLEAIKEDGREDVMEFYGLIFSTENL 299
 DB 1696 -----TMKEIISYHQF--FYKLIDPSDQL 1719

RESULT 4

apolipoprotein B - pig (fragment)

C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 13-Aug-1999
 C:Accession: I46569
 R:Puttelli, C.; Maeda, N.; Ebert, D.L.; Kaiser, M.; Lund-Katz, S.; Sturley, S.L.; Kodo

J. Lipid Res. 34, 1323-1335, 1993
 A:Title: Nucleotide sequence encoding the carboxyl-terminal half of apolipoprotein B
 A:Reference number: I46569; MUID:94014802; PMID:8409766
 A:Accession: I46569
 A:Status: preliminary; translated from GB/EMBL/DBDUB
 A:Molecule type: DNA
 A:Residues: 1-2629 <PDR>
 A:Cross-references: GB:L11235; NID:g164371; PIDN:AAA74655.1; PID:g951375
 C:Genetics:
 A:Gene: APOB
 A:Introns: 1984/1; 2022/2; 2083/3
 C:Superfamily: apolipoprotein B

Query Match 5.6%; Score 99.5; DB 2; Length 2629;
 Best Local Similarity 18.7%; Pred. No. 15;
 Matches 72; Conservative 61; Mismatches 140; Indels 113; Gaps 16;

QY 36 RFOLEP-----GSRIC-----LY-----EDGTELEDYFPSPVNAEL----- 69
 DB 2255 RFOLEPGRANNTGDELCKMMVTEVGEVLQIYSKIHSGLEILSYFODLMERSKLNARK 2314
 QY 70 -----VLTLGAMOGYSDIRRFSAFHEPOVGLIQAQQLC----- 108
 DB 2315 IKTFPDSKYQVLTVDVSEYGRQLKSLSDVQALSDLSINITTELSLQIFLECFIOEI 2374
 QY 109 -DEQAPORQLADLLHNVSQNIAMETRAEDPWFEGLESRFQSGYRSCSRI-- 164
 DB 2375 EEEFLRLKEKRLADPINDIQNINNTFTVAPLGRRLKENLDSFGMLNFIQNTLMEA 2434
 QY 165 RSYLREVSST-----PSYVG-----AEQEEFLRYLQSGMQRRLRSMQYNGSYFDR 209
 DB 2435 SOELQOLHQYIKALRKEFYDPSMGMVTKYYLEKEKVINLTKNLVDYK--DFHSKYTVS 2492
 QY 210 GAKGSRCTPEGWFCGPFDMDSCLS-----RHSINPYSNRESRIIFSTWNL 260
 DB 2493 ATDFASQLSQVEQVQVQ--DIQRYLSITLADADGKGKELAEISSRAQETI--KSAV-- 2546
 QY 261 IIEKKRTIIPVLEAIKEDGREDVMEFYGLIFSTENLKVHYCHK-- 308
 DB 2547 -----TMKEIISYHQF--FYKLIDPSDQLSDYK--KFLAEFLRLDLSIQKYNM 2594

```

Oy 309 -----KTHKLNCDPSRIYKPOTRL 328
Db 2595 ELRYMKLELOSDTVMDMRPIKV 2620

RESULT 5
A25074
vimentin - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence revision 10-May-1996 #text_change 13-Aug-1999
C:Accession: S13115; A25074; J02034; S08487; I54056
R:Honore, B.; Madsen, P.; Basse, B.; Andersen, A.; Walbum, E.; Celis, J.E.; Leffers, H.
Nucleic Acids Res. 18, 6692, 1990
A:Title: Nucleotide sequence of cDNA coding the complete coding part of the human vime
A:Reference number: S13115; M01D:91067467; PMID:2251132
A:Accession: S13115
Molecule type: mRNA
A:Residues: 1-466 <HON>
A:Cross-references: EMBL:X56134; NID:937849; P1DN:CA39600.1; P1D:937850
R:Ferrari, S.; Beldini, K.; Macnatek, L.; Rittling, S.; Calabretta, B.; De Riel, J.K.;
Mol. Cell. Biol. 6, 3614-3620, 1986
A:Title: Coding sequence and growth regulation of the human vimentin gene.
A:Reference number: A25074; M01D:87089701; PMID:3467175
A:Accession: A25074
Molecule type: DNA
A:Residues: 1-41, 'D', '43-441', 'F', '443-466 <FER>
A:Cross-references: GB:M14144; NID:9340218; P1DN:AAA61279.1; P1D:9340219
R:Perréau, J.; Lillienbaum, A.; Vasseur, M.; Paulin, D.
Gene 62, 7-16, 1988
A:Title: Nucleotide sequence of the human vimentin gene and regulation of its transcript
A:Reference number: J02034; M01D:88226018; PMID:3371665
A:Accession: J02034
Molecule type: DNA
A:Residues: 113-200, 'S', '202-264', 'S', '266-277', 'I', '279-466 <PER>
A:Note: the authors translated the codon GCG for residue 287 as ASP
R:Sommers, C.L.; Walker-Jones, D.; Heckford, S.E.; Worland, P.; Valverius, E.; Clark, R.
Cancer Res. 49, 4258-4263, 1989
A:Title: Vimentin rather than keratin expression in some hormone-independent breast can
A:Reference number: S08487; M01D:89303836; PMID:2472886
A:Accession: S08487
A:Status: preliminary
Molecule type: mRNA
A:Residues: 1-113, 'P', '114-135 <SOM>
R:Gupta, A.K.; Aubin, J.E.; Wayer, M.M.
Gene 86, 303-304, 1990
A:Title: Isolation of a human vimentin cDNA with a long 3'-noncoding region from a human
A:Reference number: I54056; M01D:90215314; PMID:2323579
A:Accession: I54056
A:Status: preliminary; translated from GB/EMBL/DBJ
Molecule type: mRNA
A:Residues: 167-466 <RES>
A:Cross-references: GB:M25246; NID:9340233; P1DN:AAA61282.1; P1D:9340234
C:Genetics:
A:Gene: GDB:VIM
A:Cross-references: GDB:119630; OMIM:193060
A:Map position: 10p13-10p13
A:Introns: 188/2; 208/3; 294/3; 336/3; 410/2; 425/1; 453/3
C:Superfamily: cytoskeletal keratin
C:Keywords: coiled coil; intermediate filament; phosphoprotein

Query Match 5.5%; Score 99; DB 2; Length 466;
Best Local Similarity 20.8%; Pred. No. 1.7;
Matches 65; Conservative 47; Mismatches 111; Indels 90; Gaps 15;

Oy 27 OEYLKRGCLRPOLPGRSGRLCYEDGTELPEDYPPSPVDAEYLITLTCGAMOGVYSIR 86
Db 180 EDIMR---LREKIGEE---MIGREAEMLTOSRQDVNNAISARLDLREKVESIQEET- 231
Oy 87 RLSAFHEPOVGLTQAA-----QQLLDEQAPORQRLADLLHNSQNTAATRAEDPPW 141
Db 232 APLAKRIHEIEIQLAQLOEOHVDLDVDSKPLDITLAALRDVQRQT-ESVAANKNLOEAREW 290

RESULT 6
A43803
vimentin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 12-Feb-1993 #sequence revision 12-Feb-1993 #text_change 20-Jun-2000
C:Accession: A43803; S32832; J50241; S12774; S14526; A31951; B31951; S32627
R:Capetanaki, Y.; Kuisk, I.; Rothblum, K.; Starnes, S.
Oncogene 5, 645-655, 1990
A:Title: Mouse vimentin: structural relationship to fos, jun, CREB and tpr.
A:Reference number: A43803; M01D:90265604; PMID:2140597
A:Accession: A43803
Molecule type: mRNA
A:Residues: 1-466 <CAP>
A:Cross-references: EMBL:X51438; NID:955290; P1DN:CA35803.1; P1D:955291
R:Rauscher, A.
submitted to the EMBL Data Library, April 1993
A:Description: Upstream region of the mouse vimentin gene.
A:Reference number: S32627
A:Accession: S32832
A:Status: preliminary
Molecule type: DNA
A:Residues: 1-70 <RAU>
A:Cross-references: EMBL:Z22526; NID:91518339; P1DN:CA480251.1; P1D:9293950
R:Wood, L.; Thériault, N.; Vogel, G.
Gene 76, 171-175, 1989
A:Title: Vimentin cDNA clones covering the complete intermediate-filament protein are
A:Reference number: J50241; M01D:89306653; PMID:2744479
A:Accession: J50241
Molecule type: mRNA
A:Residues: 1-155, 'EL', '158-337', 'E', '339-373', 'D', '375-466 <MOO>
A:Cross-references: GB:M24849; NID:9202367; P1DN:AAA40555.1; P1D:9202368
R:Hennekkes, H.; Kuehn, S.; Traub, P.
Mol. Gen. Genet. 221, 33-36, 1990
A:Title: Coding sequence and flanking regions of the mouse vimentin gene.
A:Reference number: S12774; M01D:90220517; PMID:2325630
A:Accession: S12774
A:Status: translation not shown
Molecule type: mRNA
A:Residues: 1-155, 'EL', '158-337', 'E', '339-466 <HEN>
A:Cross-references: EMBL:M26251; NID:9202369; P1DN:AAA40556.1; P1D:9202370
R:Podolin, P.L.; Pryszowsky, M.B.
submitted to the EMBL Data Library, October 1990
A:Description: Nucleotide sequence of a murine vimentin cDNA.
A:Reference number: S14526
A:Accession: S14526
Molecule type: mRNA
A:Residues: 1-155, 'EL', '158-163', 'E', '165-337', 'E', '339-466 <POD>
A:Cross-references: EMBL:X56397; NID:954407; P1DN:CA39807.1; P1D:954408
R:ando, S.; Tanabe, K.; Honda, Y.; Sato, C.; Inagaki, M.
Biochemistry 28, 2974-2979, 1989
A:Title: Domain- and sequence-specific phosphorylation of vimentin induces disassemb
A:Reference number: A31951; M01D:89302884; PMID:2500966
A:Accession: A31951
Molecule type: protein
A:Residues: 5-12; 14-28; 37-69 <AND>
A:Accession: B31951

```

A:Molecule type: protein
 A:Residues: 5-12;14-69 <AN2>
 C:Superfamily: cytoskeletal keratin
 C:Keywords: coiled coil; intermediate filament; phosphoprotein
 F:104-138/Region: coil 1A
 F:147-247/Region: coil 1B
 F:264-282/Region: coil 1B
 F:291-411/Region: coil 2B
 F:7,25,39,51,66/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase and F
 F:9,10,21,26,34,42/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #statu
 F:47/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status experim

Query Match 5.5%; Score 99; DB 2; Length 466;
 Best Local Similarity 20.1%; Pred. No. 1.7; Mismatches 113; Indels 90; Gaps 14;
 Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;

Db 27 QEVLRKGLRQLPBERGSRRLCYEDGETELTDEYPPVDNAELVLTGAMGVSDIR 86
 180 EDIMR---LRKLOEE---MLQREASTIQSFROVDNASTARLDLKERVESIQEET- 231
 87 RFLSAFHEPOVGLIOAA-----QQLLDEQAPORORLLADLLHNVSONIAETRAEDPPW 141
 232 AFLKRLHDEETIOELQAOIOEQHVIDVDSKPDLTALRDVROQY-ESVAKNLQEAEEW 290

Qy 142 F-----EGLESRFQSKSGYLK-----YSCESRIR---S 166
 Db 291 YKSKFADLSEANRNNDLRQAKQESNRYRQVOSTLCEVDALKTNYSLERQKMEEN 350
 Qy 167 YLREVSSYPTVGAEOEFLRVLSMCQRLSRMQYNSYPRGAKGSGRLCTPEGWFC 226
 351 FALSAANYODTIG-RLQDEIQNMKEEMARHLREYQ-----DLLNV 389.
 Db 227 QGPFDMDSCLSRHSINPYSNRESRI-----LFTWNLDHT-----IEKRTIIPTL 272
 390 KMALDIEIATYRKLL---EGEESRISLPLPTFSSILNRETMLSESLPLVDTHSKRTL--L 443
 Qy 273 VEAIKEDDGREVD 285
 Db 444 IKTVETRDGOVIN 456

RESULT 7
 AF0627
 hellicase IV (75 kD hellicase) [imported] - Salmonella enterica subsp. enterica serovar Ty
 C:Species: Salmonella enterica subsp. enterica serovar Typh
 A:Note: This species has also been called Salmonella Typh
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AF0627
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connetton, P.; Croplin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moulé, S.; O'Garra, P.
 F:413, 848-852, 2001
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-684 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08201.1; PID:G16502250; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1097
 C:Superfamily: hellicase IV

Query Match 5.5%; Score 98.5; DB 2; Length 684;
 Best Local Similarity 22.9%; Pred. No. 3.1; Mismatches 148; Indels 99; Gaps 22;
 Matches 89; Conservative 52; Mismatches 148; Indels 99; Gaps 22;

Qy 26 QEVLRKGLRQLPBERGSRRL---CLYEDG-TELTEDYFPSPVD-----NAEL 69
 152 CREITWRK-CLAWLQDSESGRQHQHQAADAMLEAHADFQIIESSPLNPQARAVNGES 210
 Db 70 VLITLGGAMGVSDI---RFLSAFHEPOVGLI-----QAAQLLDEQAPOR---Q 116

Db 211 SLVLVAGSGKTSVLVARACWLLARQADAGQILLAFGRKAAEEM--DRIRERLATE 268
 Qy 117 RLADLLHNVSONIAETRAEDPPWFEGLES-----RFQSKGYLRVYSCESR-----IRS 166
 269 EITARTFSLALYIT-IOGSKKAPVSKLESDATARRHQLFHTNRQOCSEKKAQAKGRO 327
 Db 167 YLREVSSYPTVGAEOEFLRVLSMCQRLSRMQYNSYPRGAKGSGRLCTPEGWFC 224
 328 WLEEMQWVVEGNEFWDEETLQRRILAPLRDWRVSLMRHG-----GAQAEIMAGAPE--- 379
 Qy 225 SCQGPFDMDSCLSRHSINPYSN---RESRIILFTWNLDHT-----IEKRTIIP----- 270
 380 ECRLEFGKRITLAPLPLAMKSALENAVDPS-GLTHQAMVILLEGRTISPMKHLIYD 437
 Db 271 -----TLVEAIKEDDGREV-----DMEYFGILFTSEMLKLVHY----- 305
 438 EFQDISPQRAALLERLKRQNSQTLLFANGDDMOAIT--RFGAQLSLTFAHQTFEGEGH 495
 Qy 306 CHKTKTKLNCDPSPRIYKPTRLRKRP 333
 496 CHLDYTRFN---SRIGDIANRFVQNP 520

RESULT 8
 I48128
 vimentin - Chinese hamster (fragment)
 C:Species: Crictetus griseus (Chinese hamster)
 C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 13-Aug-1999
 C:Accession: I48128
 R:Blomendaal, H.; Quax, W.; Quax-Jeuken, Y.; Dodeмонт, H.; Ramaekers, F.; Dunia, I.;
 Mol. Biol. Rep. 9, 115-118, 1983
 A:Title: Organization and expression of the vimentin gene.
 A:Reference number: I48128; MUID:83297272; PMID:6688458
 A:Accession: I48128
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-448 <RES>
 A:Cross-references: GB:M16718; NID:G191232; PIDN:AAA37029.1; PID:G387058
 C:Superfamily: cytoskeletal keratin

Query Match 5.5%; Score 98; DB 2; Length 448;
 Best Local Similarity 20.1%; Pred. No. 2; Mismatches 113; Indels 90; Gaps 14;
 Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;

Qy 27 QEVLRKGLRQLPBERGSRRLCYEDGETELTDEYPPVDNAELVLTGAMGVSDIR 86
 162 EDIIR---LRKLOEE---MLQREASTIQSFROVDNASTARLDLKERVESIQEET- 213
 87 RFLSAFHEPOVGLIOAA-----QQLLDEQAPORORLLADLLHNVSONIAETRAEDPPW 141
 214 AFLKRLHDEETIOELQAOIOEQHVIDVDSKPDLTALRDVROQY-ESVAKNLQEAEEW 272
 Qy 142 F-----EGLESRFQSKSGYLK-----YSCESRIR---S 166
 Db 291 YKSKFADLSEANRNNDLRQAKQESNRYRQVOSTLCEVDALKTNYSLERQKMEEN 332
 Qy 167 YLREVSSYPTVGAEOEFLRVLSMCQRLSRMQYNSYPRGAKGSGRLCTPEGWFC 226
 333 FALSAANYODTIG-RLQDEIQNMKEEMARHLREYQ-----DLLNV 371
 Db 227 QGPFDMDSCLSRHSINPYSNRESRI-----LFTWNLDHT-----IEKRTIIPTL 272
 372 KMALDIEIATYRKLL---EGEESRISLPLPTFSSILNRETMLSESLPLVDTHSKRTL--L 425
 Qy 273 VEAIKEDDGREVD 285
 Db 426 IKTVETRDGOVIN 438

RESULT 9
 VEHY
 vimentin - golden hamster


```

C:Species: Mesocricetus auratus (golden hamster)
C>Date: 19-Feb-1984 #sequence_revision 27-Nov-1985 #text_change 22-Jun-1999
C:Accession: A90842; A93953; A39731; A02959
R:Quax, W.; Egberts, W.V.; Hendriks, W.; Quax-Juken, Y.; Bloemendal, H.
Cell 35, 215-223, 1983
A:Title: The structure of the vimentin gene.
A:Reference number: A90842; MUID:84026520; PMID:6194898
A:Accession: A90842
A:Molecule type: DNA
A:Residues: 1-464 <OU>
A:Cross-references: CB:K00927; NID:g191465; PIDN:AA37104.1; PID:g387077
R:Quax-Juken, Y.E.F.M.; Quax, W.J.; Bloemendal, H.
Proc. Natl. Acad. Sci. U.S.A. 80, 3548-3552, 1983
A:Title: Primary and secondary structure of hamster vimentin predicted from the nucleotide
A:Reference number: A93953; MUID:85221633; PMID:6304716
A:Accession: A93953
A:Molecule type: mRNA
A:Residues: 1818EPG, 25-40, 'A', 42-113, 'D', 115-180, 'T', 182-464 <OU>
A:Experimental source: lens
R:Chou, Y.H.; Ngai, K.L.; Goldman, R.
J. Biol. Chem. 266, 7335-7328, 1991
A:Title: The regulation of intermediate filament reorganization in mitosis. p34(cdc2) ph
A:Reference number: A39731; MUID:91210232; PMID:2019567
A:Accession: A39731
A:Molecule type: protein
A:Residues: 36-41, 'A', 42-47, 'T', 48-49, 50-62, 63-67 <CHO>
A:Note: the phosphorylated residue described as 40-Ser may be 41-Ser; nucleic acid sequence
C:Comment: The initiator Met is not shown.
C:Comment: Vimentin occurs in intermediate-sized filaments in various nonepithelial cells.
C:Genetics:
A:Introns: 187/2, 207/3; 239/3; 293/3; 335/3; 409/2; 424/1; 452/3
C:Superfamily: cytoskeletal keratin
C:Keywords: acetylated amino end; coiled coil; intermediate filament; phosphoprotein
F:1-464/Domain: tail <TE>
F:1/Modified site: acetylated amino end (Ser) #status predicted
F:4/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status predicted
F:52,64/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status experimental

Query Match 5.5% Score 98; DB 1; Length 464;
Best Local Similarity 20.1%; Pred. No. 2;
Matches 63; Conservative 47; Mismatches 113; Indels 90; Gaps 14;

OY 27 QEVLRGGLRFPGRSRLCYEDGTELEDFYPPSPDPAELVLTGLQAMQGYSDIR 86
178 EDLMR--LRKRLQEE--MLQREASTLQSPRDVDNASTARLADLRKVESLQEEI- 229
DB 87 RLTSAPHEPQVLIQA-----QQLLDEQAPQQRILLADLLHNSVNTALETADPPW 141
230 AFLKRLHDEIQLDAQIQEQHYIDVDVSKPDLTALRDVROOY-ESVAANKLDEAEW 288
OY 142 F-----EGLSERFQSKSGYLRY-----SCESRIR-----S 166
DB 289 YKSKFADLSEANRRNNDLRLQAKQESNEYRQVOYLCEVDALKTGNETSIFEROMRMEEN 348
OY 167 YLREVSYSYVGAQAOEELVLQSMQCRQYMGYSYFGDAGKGSRLCTPREGWFC 226
DB 349 FALDAANYODTIG-RLQDEIQMKKEAARHLREYO-----DLNLY 387
OY 227 QGFQDMDSCLSHSINPYSNRSLT-----FSWNTDHL-----LEKKRTIPL 272
DB 388 KMALDIEIATYRKLL--EGESRSILPLPFFSSLNLRFTNNESSLVLYDTHSKRL--L 441
OY 273 VEAKRQDGRVD 285
DB 442 IKTVERDRDQVIV 454

RESULT 10
522119
vimentin - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 13-Aug-1999
C:Accession: 522119; J01399

```

```

R:Bussenaekers,M.J.G.; Verhaegh,G.C.W.; van Broeken,A.; Debruyne,F.M.J.; Schalken,  
submitted to the EMBL Data Library, October 1991  
A:Description: differential expressions of vimentin in rat prostatic tumors: complete  
A:Reference number: S22119  
A:Accession: S22119  
A:Molecule type: mRNA  
A:Residues: 1-466 <DS>  
C:Cross-references: EMBL:X62952; NID:g57479; PIDN:CMA4722.1; PID:g57480  
A:Experimental source: prostatic tumor; cell line Dunning R-3327, strain Fischer Cope  
R:Bussenaekers,M.J.G.; Verhaegh,G.W.C.T.; Van Bokhoven,A.; Debruyne,F.M.J.; Schaik  
Biochem. Biophys. Res. Commun. 182, 1254-1259, 1992  
A>Title: Differential expression of vimentin in rat prostatic tumors.  
A:Reference number: JQ1389; MUID:92171936; PMID:1540169  
A:Accession: JQ1389  
A:Molecule type: mRNA  
A:Residues: 1-466 <BUW>  
C:Cross-references: EMBL:X62953  
A:Experimental source: prostatic tumor; cell line Dunning R-3327; strain Fischer Cope  
C:Superfamily: cytoskeletal keratin  
C:Keywords: phosphoprotein
```

```
Query Match          5.5%; Score 98; DB 2; Length 466;  
Best Local Similarity 20.1%; Pred.No.2.1;  
Matches      63; Conservative   47; Mismatches 113; Indels    90; Gaps     14;
```

```
OY  27 OEVLKRCGRPLPFRSGRLCLVEDGETLFEDFPSPDVAEYLTLTGAAOGVYSIR  86  
Db   ::::| | | | | | | | | | | | | | | | | | | | | | | | | |  
180 EDIMR-----LRKLGE-----MLGREPAESTLOSPOVDNAIARLDERKVESLOEBI - 231  
OY  87 RFISANHEPGVALIA-----QQLTDEQAQRQLADLIHNSONTAAETRADEPPW  141  
Db   ::::| | | | | | | | | | | | | | | | | | | | | | | | | |  
232 AFLKRKHDELIELAQIQBOHQVIDVSKSPDLTALNDVRQY-EVSAAKNQDAEW  290  
OY  142 F-----EGLESRPQSKGYLY-----SCSERIR----S  166  
Db   ::::| | | | | | | | | | | | | | | | | | | | | | | | | |  
291 YKKRFADLEARNRNNDALROAKQENSDRYRQVSLTCVDAQLKGTNESLERQMEEHN  350  
OY  167 VYLEVSSYSTYGAEAGEFLFAYLGSMCKRLRSMGNGSTFDGAGKGSRLLCTPFGMTSC  226  
Db   351 FALEAAANDDTIG-RIOEIOMKEEKARKRIREQ-----DLLNV  389  
OY  227 QGFDMDCLSHSHSNPSNRERTL-----FSTWNIDHI-----TEKRTTIPTL  272  
Db   ::::| | | | | | | | | | | | | | | | | | | | | | | | | |  
390 KHALDLKITVKRL--EGBSRILPLPFSSLNIRETNESLPDYTHSKRTL--L  443  
OY  273 VEAIKRGDGREVD  285  
Db   444 IKTVETRDGOVIN  456
```

```
RESULT 11  
148771  
SI(PW7) - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999  
C:Accession: I48771  
R:Hemenway,C.; Kalif,M.; Stravenagen,J.; Walthall,D.; Robbins,D.  
Nucleic Acids Res. 14, 2539-2554, 1986  
A>Title: Sequence comparison of alleles of the fourth component of complement (C4) an  
A:Reference number: I48274; MUID:86176748; PMID:3008092  
A:Accession: I48771  
A>Status: preliminary; translated from GB/EMBL/DDBI  
A:Molecule type: mRNA  
A:Residues: 1-594 <RES>  
C:Cross-references: EMBL:X06454; NID:g54105; PIDN:CAA29760.1; PID:g54106  
C:Superfamily: alpha-2-macroglobulin
```

```
Query Match          5.5%; Score 98; DB 2; Length 594;  
Best Local Similarity 24.1%; Pred.No.2.8;  
Matches      56; Conservative   29; Mismatches  82; Indels    72; Gaps     12;
```

```
18 KFCVAGRSCDEV-LRKG-CILRFOLPERGLCY---EDGTETLEDYPFSVPDNEL  69  
:|:::| | | | | | | | | | | | | | | | | | | | | | | | | | |  
:
```

Db 327 KLGLSGMAIDITLLSGFHALRGDLKLTSLSDRYVSHFETDGPVILL-YFDSVPTTREC 385
 70 VLLTLGAMOGYVSDIRRFSAFHEPOVGLIOAQQLLCD-----EQAPORQL 118
 Db 386 V-----GF-----GASGEVYVGLVQPSAVLYDYSPDHKCSVYAPTKSQL 428
 Oy 119 LADLHNVSQNIATETAEPPWFEGLESFPOKSGY-LRYSC-----160
 429 LATL---CSGDVCCAGCKPPLRLSLERVEDKGRMRACVYPRVEGYTKVLRED 485
 Oy 161 -----ESRISYLEVSSYPTVGAEQAEFLRLVLSMORLSMOYNGSVFEDGAK 212
 486 GRAAFRLPESKITQVLFPRIDTMASIG-----QTRNLSRTSKRL-LEPKETYLIMGMD 539
 Oy 213 G 213
 Db 540 G 540

RESULT 12

Agrobacterium tumefaciens (strain C58, Dupont) plasmid
 Accession: AG3174
 Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
 C:Accession: AG3174
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutlyavin, T.; Levy, R.; Li, M.; McCelli
 science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AG3174
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-571 <KUR>
 A:Cross-references: GB:AE008687; PIDN:AA145813.1; PID:g17743551; GSPDB:GN00188
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: atrc
 A:Genome: plasmid
 C:Superfamily: acetolactate synthase large chain; thiamin pyrophosphate-binding domain

Query Match 5.48; Score 96.5; DB 2; Length 571;
 Best Local Similarity 21.6%; Pred. No. 3.6;
 Matches 79; Conservative 49; Mismatches 122; Indels 115; Gaps 20;
 Oy 10 LRALRSRKRKGVAGRSQVLRKGLRQFPERGSRCLYEDGELTEDYFPSPDMAEL 69
 259 LASLVGKSL--GRHVEIVREADLVLLGTFTN-----QNGTDNRQYSP---DAOI 306
 Db 70 VLLTLGAMOGYVSDIRRFSAFHEPOVGLIOAQQLLCDQAPORQL-LADL-LHNVS 127
 307 IHDIDTQDEQGRNFARLVGDARETLAQL-----RRIRLCIDLHTVS 351
 Oy 128 QNIAETRAEDPWFEGLESFPOKSGYLYRSCS-----RISYLEVSSYPTVGAE 181
 352 REAVCNRLRES--WRKFNDR---RGY--YSSEASPLRPERIWAELQGVIDENTITYAD 403
 Db 182 AOEFLFVLGSM--CORLSMOYNGSVFEDGAKG-----GSLRLCPE-----221
 404 ASYSSMNVLLQRLISSRTVLT-----PRGLAGLVGVPALGAIVAAVESDVAVLG 457
 Oy 222 -GWFSCGPPDMSCLSRHSINPYSNRESILFTW--NLHIIIEKKRTIIPTEALIK-- 277
 458 DGGFA-----HS-----MALEFLIV--RMQIPITIVVANG 486
 Db 278 -----EODGREVDEYEGLLFTSENLKLVHIVCHKRTTHKLNCDPSRIYKQTRLRKOP 333
 487 ILGFORDAETVYKFGKTYTACHFAE-----VDHVKAIAHACGDAVAVTQPODKAHMOK 539
 Oy 334 VKRKQ 338

Db 540 VKRK 544

RESULT 13

COI intron 10 protein - Podospira anserina mitochondrion
 C38888
 C:Species: mitochondrion Podospira anserina
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 03-Aug-2001
 C:Accession: C38888
 R:Cummings, D.J.; Michel, F.; McNally, K.L.
 Curr. Genet. 16, 381-406, 1989
 A>Title: DNA sequence analysis of the 24.5 kilobase pair cytochrome oxidase subunit I
 A:Reference number: A48327; MUID:90124722; PMID:2558809
 A:Accession: C38888
 A:Molecule type: DNA
 A:Status: preliminary
 A:Residues: 1-397 <CUM>
 A:Cross-references: GB:X55026; GB:M30937; GB:M61734; MID:g14030; PIDN:CAA38788.1; PID
 C:Genetics:
 A:Genome: mitochondrion
 C:Superfamily: Podospira anserina mitochondrion COI intron 10 protein; COI intron 9 p
 C:Keywords: mitochondrion
 F:157-380/Domain: COI intron 9 protein homology <C19>

Query Match 5.28; Score 93.5; DB 2; Length 397;
 Best Local Similarity 22.4%; Pred. No. 4;
 Matches 64; Conservative 39; Mismatches 112; Indels 71; Gaps 12;

Oy 45 RLCLYEDGTELT---EDYFSPVDMALVLLTLGAMOGYVSDIRRFSAFHEPOVGLI 100
 49 KILVIMNPQITTKARSNKYKGFTEFLGLCA-----VWGISSEAIRLLPISLKRIDL 102
 Db 103 FALLOYSIRLEKGYFLMNKLKYSTLASDYSENSLANRNKNSPFKEKEDFEMWLAGIT 162
 Oy 151 -----SKSGYLYRSCESRIRS---YLRVSSYPTVGAEQAEFLR-----188
 163 DGDGCFQVSKKGVASLEIVQLRDKRLIYLK-QYGGAVKALHGDVYLRRLHKKAGL 221
 Db 189 -----VLSMCQRLRSQYNG--SYPRGAKGSRICTPGWSGCGPPDMSCLSRHSI 241
 222 SLINGINGLIRNPIRIQLGRICMLYDIKLDQPLTYMGWLS--GFFPDGSI-----274
 Oy 242 NPYSNRESRLFSFWMNDHIEKKRTIIPTE-----AIKROD 280
 Db 275 --YLNKSGQIFITAS-----QKNRFLDALVELYGGTYAMVAKOD 313

RESULT 14

GNLIGS
 HIV-1 retropepsin (EC 3.4.23.16) - simian immunodeficiency virus (African green monkey
 N:Contans: endonuclease (EC 3.1.1.-) retropepsin (EC 3.4.23.16); RNA-directed DNA P
 C:Species: simian immunodeficiency virus, HIV
 C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 03-Jun-2002
 C:Accession: B28873
 R:Franchini, G.; Guigo, C.; Guo, H.G.; Gallo, R.C.; Collalti, E.; Fargnoli, K.A.; Hal
 Nature 328, 539-543, 1987
 A>Title: Sequence of simian immunodeficiency virus and its relationship to the human
 A:Reference number: A28873; MUID:87287229; PMID:3457350
 A:Accession: B28873
 A:Molecule type: DNA
 A:Residues: 1-1054 <FRA>
 A:Cross-references: EMBL:M19499
 C:Comment: Specific enzymatic cleavages may yield mature proteins including protease,
 C:Genetics:
 A:Gene: pol
 C:Superfamily: pol polypeptide
 C:Keywords: aspartic proteinase; hydrolase; nucleotidyltransferase; polypeptide; reye
 F:106-204/Product: retropepsin #status predicted <RTP>
 F:130/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 5.28; Score 92.5; DB 1; Length 1054;
 Best Local Similarity 22.88; Pred. No. 18;
 Matches 56; Conservative 34; Mismatches 93; Indels 63; Gaps 11;

13 LKSPKFGVAGRSQCEVLRKGCRLRFLPERGSLCLYEDGETLEDYFPPSPDAAE----68
 Db IRIPHPGLARKRRITVLIDIGDAVFSP-----LDEFFROYTFITLPSV--NNAEFGKR 347
 296 IRLPGLARKRRITVLIDIGDAVFSP-----LDEFFROYTFITLPSV--NNAEFGKR 347
 69 LVLTILGQAMOGC---VSDIRRLSAPHE--POVGLQAQOQL--CDEQAPQRRL--118
 Oy 348 YIKYKLPQGMKSPALFOYTMRAVLEPERKANPDVTLVQYMDILIASDRDLEHDKRVL 407
 119 -LADLHNVSQNIARADPPM--FEGLE-----SRQSSSG 155
 408 QLKELNLSIGSTPEKFKQDPFPQMMGYELPTKMKLOKIELPQRETWYNQIKRYGV 467
 156 LKYSCE-----SRISYLEVSSYPTVGAEOEPLNVLGSMCRSLSMQYNG 204
 Db 468 LKMAQIYIPGIRKTHLCRLIRNGKMTLEAVONTKMAELEYENNI-----ILSQDEG 520
 Oy 205 SYFDNG 210
 Db 521 CYDEG 526

RESULT 15
 S54784
 sex-limited protein SIP(w7) - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 27-Oct-1995 #sequence.revision 03-Nov-1995 #text.change 16-Jul-1999
 C:Accession: S54784; 148770
 R:Ogata, R.T.; Zepf, N.E.
 J. Immunol. 147, 2756-2763, 1991
 A:Title: The murine SIP gene. Additional evidence that sex-limited protein has no biologic
 A:Reference number: S54784; MUID:92013090; PMID:1918990
 A:Accession: S54784
 A:Status: preliminary; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1735 <OGA>
 A:Cross-references: EMBL:M64933; NID:9200988; PIDD:AAA0117.1; PTD:9200989
 R:Hemenway, C.; Kalf, M.; Stavenhagen, J.; Walthall, D.; Robins, D.
 Nucleic Acids Res. 14, 2539-2554, 1986
 A:Title: Sequence comparison of alleles of the fourth component of complement (C4) and s
 A:Reference number: 148274; MUID:6116748; PMID:3008092
 A:Accession: 148770
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 634-641, 'L', 643-828 <RES>
 A:Cross-references: EMBL:X06455; NID:954102; PIDD:CAA29761.1; PTD:9899336
 C:Genetics:
 A:Insertions: 22/2; 86/3; 154/1; 177/3; 207/2; 235/1; 267/2; 302/3; 347/1; 385/3; 444/3; 50
 3; 1219/1; 1296/3; 1352/3; 1372/3; 1404/1; 1464/2; 1494/3; 1519/3; 1554/1; 1584/1; 1617/
 C:Superfamily: alpha-2-macroglobulin

Query Match 5.18; Score 92; DB 2; Length 1735;
 Best Local Similarity 27.68; Pred. No. 38;
 Matches 50; Conservative 20; Mismatches 63; Indels 48; Gaps 11;

18 KGVAGRSQCEV--LRKG--CLRFQLPERSRLCLY-----EDGETLEDYFPPSPDAAEL 69
 Db 1468 KGLSGMAIADITLLSGFHALRADLEKITSLSDRYSHRETDGPHVLL--YDSDVPTTREC 1526
 Oy 70 VLLTILGQAMOGYVSDIRRLSAPHEPOVGLQAQOQLD-----EQAPQRRL 118
 Db 1527 V-----GF-----GASQEVVGLVOPASAVLYLDIYSPDHKCSVYAAVPTSQL 1569
 Oy 119 LADLHNVSQNIARADPPM--FEGLE-----SRQSSSG 155
 Db 1570 LKYL---CSGDVCCACAGKCPRLRLSLRERVEDKQDYRRFAC-----YYHOV--EYGET 1619
 Oy 178 V 178

Db 1620 V 1620

RESULT 16
 A29176
 sex-limited protein precursor - mouse
 N:Alternate names: complement C4-related sex-limited protein
 C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1988 #sequence.revision 30-Jun-1988 #text.change 16-Jul-1999
 C:Accession: A29176; B21692; A41195; B29059; B60227; 179467
 R:Ogata, R.T.; Seplich, D.S.
 J. Immunol. 135, 4239-4244, 1985
 A:Title: Murine sex-limited protein: complete cDNA sequence and comparison with murin
 A:Reference number: A29176; MUID:86060918; PMID:3840826
 A:Accession: A29176
 A:Molecule type: mRNA
 A:Residues: 1-1736 <OGA>
 R:Nonaka, M.; Takahashi, M.; Natsume-Sakal, S.; Nonaka, M.; Tanaka, S.; Shimizu, A.;
 Proc. Natl. Acad. Sci. U.S.A. 81, 6822-6826, 1984
 A:Title: Isolation of cDNA clones specifying the fourth component of mouse complement
 A:Reference number: A94013; MUID:85038607; PMID:6208559
 A:Accession: B21692
 A:Molecule type: mRNA
 A:Residues: 651-749, 'H', 751-774, 'D', 776-802, 921-1040 <NON>
 R:Ogata, R.T.; Seplich, D.S.
 Proc. Natl. Acad. Sci. U.S.A. 81, 4908-4911, 1984
 A:Title: Genes for murine fourth complement component (C4) and sex-limited protein (S
 A:Reference number: A41195; MUID:84272739; PMID:6589636
 A:Accession: A41195
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1345-1438, 'R', 1440-1544 <OG2>
 A:Cross-references: GB:K02293; NID:9199611; PIDD:AAA39662.1; PTD:9199612
 R:Tooi, M.; Levi-Strauss, M.; Duponchel, C.; Meo, T.
 Pflörs, Trans. R. Soc. Lond. 306, 389-394, 1984
 A:Title: Sequence heterogeneity of murine complementary DNA clones related to the C4
 A:Reference number: A93753
 A:Accession: B29059
 A:Molecule type: mRNA
 A:Residues: 1355-1335, 'G', 1337-1373 <TOS>
 R:Ogata, R.T.; Zepf, N.E.
 Eur. J. Immunol. 20, 1607-1610, 1990
 A:Title: C4 from C4-High and C4-Low mouse strains have identical sequences in the reg
 A:Reference number: A60227; MUID:90353398; PMID:2387317
 A:Accession: B60227
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1096-1139 <OG3>
 A:Cross-references: GB:X55495
 R:Nonaka, M.; Kimura, H.; Yeul, Y.D.; Yokoyama, S.; Nakayama, K.; Takahashi, M.
 Proc. Natl. Acad. Sci. U.S.A. 83, 7883-7887, 1986
 A:Title: Identification of the 5'-flanking regulatory region responsible for the diff
 A:Reference number: 159084; MUID:87017050; PMID:3464002
 A:Accession: 179467
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-21 <RES>
 A:Cross-references: GB:M14226; NID:9199615; PIDD:AAA99684.1; PTD:9554239
 C:Genetics:
 A:Superfamily: alpha-2-macroglobulin
 F:1-19/Domain: signal sequence #status predicted <SIG>
 F:20-1736/Product: sex-limited protein #status predicted <MAT>

Query Match 5.18; Score 92; DB 2; Length 1736;
 Best Local Similarity 27.68; Pred. No. 38;
 Matches 50; Conservative 20; Mismatches 63; Indels 48; Gaps 11;

18 KGVAGRSQCEV--LRKG--CLRFQLPERSRLCLY-----EDGETLEDYFPPSPDAAEL 69
 Db 1468 KGLSGMAIADITLLSGFHALRADLEKITSLSDRYSHRETDGPHVLL--YDSDVPTTREC 1526
 Oy 70 VLLTILGQAMOGYVSDIRRLSAPHEPOVGLQAQOQLD-----EQAPQRRL 118

Db 1527 V-----GF-----GASQEVVGLVQPSASVLYYYSPDHKCSVFYAPPTSOL 1569

QY 119 LADLHNVSQINAEETRAEDPPMEFEGLESFPOSKSGY-LRVSCSRIRSYLREVSYSPT 177

Db 1570 LALL-----CSQDVQCAEGKCPRLRLSLERVEDKGRIMAFAC-----YTHOV-EXGFT 1619

QY 178 V 178

Db 1620 V 1620

RESULT 17

D70441

oxido/reductase iron sulfur protein - Aquifex aeolicus

C:Species: Aquifex aeolicus

C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-2000

C:Accession: D70441

R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'Neil, J.E. 1992, 353-358, 1998

File: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

Reference number: A70300; MUID:98196666; PMID:9537320

Accession: D70441

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-395 <AOE>

A:Cross-references: GB:AE000748; NID:g2983960; PIDN:AAC07510.1; PID:g2983963; GB:AE00065

A:Experimental source: strain VF5

C:Genetics:

A:Gene: glpC

C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hcf; 9-85/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match

Best Local Similarity 28.1%; Score 91.5; DB 2; Length 395;

Matches 34; Conservative 19; Mismatches 47; Indels 21; Gaps 7;

QY 229 PEPMDSCLSRHSINPYSNRSRLIFSTWINDHIEKKR-----TTFPLVEAIKEDQGRE 283

Db 206 PNDKICCGAPHLY--HONTFAPEKLEKHNKE-IEKEYDALVAVACPCCGALKEDYK- 261.

QY 284 VDEVEYF-YGLTFSENKLT-----VHIVCHKKTHKNCPSRIYKQTRKRPY 334

Db 262 -DKVVSFTETIANEDIEFGKGEKTVVHPCHYITAMKLN--PNNFYKALGKVKNAEYV 318

QY 335 R 335

Db 319 K 319

Query Match

Best Local Similarity 21.8%; Score 91.5; DB 2; Length 925;

Matches 77; Conservative 44; Mismatches 135; Indels 97; Gaps 17;

QY 28 EVLRKGCRLPOLPERGSRCLY--EDG-----TELTEDYFPSPVDMNELVLTG--GOAWOG 80

Db 513 ELKRWTSITPHRAISAHLIYFAEHDLHTHAEGLYTVGDRLLAALITRYRQDAYG 572

QY 81 YVSDI--RRFLSAFHEPQVGLIQAAQQLCDDEAPQORLLADLHNVSQINAEET----- 134

Db 573 PYRLADLRKRLKELINSQDPEDIWGLITLLEKRNQNFPIILDFLNTKNDILITVCKAL 632

QY 135 -----RAEDPPMEFEGLESRFQ-----SKSGYLRSC-----SRISYLAEVSSYST- 177

Db 633 HTSVRAHKKYCPBLKRLKQCSHNDASQYLTKTISIALDISFVLDLMTTSQLNKTSR 692

QY 178 -----VGAEOEFLNVLGSMCORLSMONGSYFPGAKGSR-----LCPEG 222

Db 693 KYAEMAGELDKKVAAPAFLOVL-----TDEGHNRCRIILAAALCKIDN 756

QY 223 WFSQCGPDMDSCLSRHSINPYSNRSRLIFSTWINDHIEKKRTIIP-----LYEAIK 277

Db 737 WL-----LKKHAYKIVKSKASKALFYSYH-GHYIQRK---YPTVNLSTLANTLN 781

QY 278 EDGREGVDWEY-FYGLL-----FTSENKLVHIVCHKKTHKLNCD 317

Db 782 SNTYAEVNFMLSLGILGSMHSGVLTIRALTSKNQK---TKQALLESLEKNC 831

RESULT 19

E81573

conserved hypothetical protein CP0465 [imported] - Chlamydia pneumoniae (strain A)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 20-Jun-2000

C:Accession: E81573

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hicke

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

A:Reference number: A81500; MUID:20150235; PMID:10684935

A:Accession: E81573

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-925 <REA>

A:Cross-references: GB:AE002208; GB:AE002161; NID:g7189387; PIDN:AAF38302.1; PID:g718

A:Experimental source: strain AK39, HL cells

C:Genetics:

A:Gene: CP0465

C:Superfamily: Chlamydia trachomatis hypothetical protein CP234

Query Match

Best Local Similarity 21.8%; Score 91.5; DB 2; Length 925;

Matches 77; Conservative 44; Mismatches 135; Indels 97; Gaps 17;

QY 28 EVLRKGCRLPOLPERGSRCLY--EDG-----TELTEDYFPSPVDMNELVLTG--GOAWOG 80

Db 513 ELKRWTSITPHRAISAHLIYFAEHDLHTHAEGLYTVGDRLLAALITRYRQDAYG 572

QY 81 YVSDI--RRFLSAFHEPQVGLIQAAQQLCDDEAPQORLLADLHNVSQINAEET----- 134

Db 573 PYRLADLRKRLKELINSQDPEDIWGLITLLEKRNQNFPIILDFLNTKNDILITVCKAL 632

QY 135 -----RAEDPPMEFEGLESRFQ-----SKSGYLRSC-----SRISYLAEVSSYST- 177

Db 633 HTSVRAHKKYCPBLKRLKQCSHNDASQYLTKTISIALDISFVLDLMTTSQLNKTSR 692

QY 178 -----VGAEOEFLNVLGSMCORLSMONGSYFPGAKGSR-----LCPEG 222

Db 693 KYAEMAGELDKKVAAPAFLOVL-----TDEGHNRCRIILAAALCKIDN 756

QY 223 WFSQCGPDMDSCLSRHSINPYSNRSRLIFSTWINDHIEKKRTIIP-----LYEAIK 277

Db 737 WL-----LKKHAYKIVKSKASKALFYSYH-GHYIQRK---YPTVNLSTLANTLN 781

QY 278 EDGREGVDWEY-FYGLL-----FTSENKLVHIVCHKKTHKLNCD 317

Matches 77; Conservative 44; Mismatches 135; Indels 97; Gaps 17;

QY 28 EVLRKGCRLPOLPERGSRCLY--EDG-----TELTEDYFPSPVDMNELVLTG--GOAWOG 80

Db 513 ELKRWTSITPHRAISAHLIYFAEHDLHTHAEGLYTVGDRLLAALITRYRQDAYG 572

QY 81 YVSDI--RRFLSAFHEPQVGLIQAAQQLCDDEAPQORLLADLHNVSQINAEET----- 134

Db 573 PYRLADLRKRLKELINSQDPEDIWGLITLLEKRNQNFPIILDFLNTKNDILITVCKAL 632

QY 135 -----RAEDPPMEFEGLESRFQ-----SKSGYLRSC-----SRISYLAEVSSYST- 177

Db 633 HTSVRAHKKYCPBLKRLKQCSHNDASQYLTKTISIALDISFVLDLMTTSQLNKTSR 692

QY 178 -----VGAEOEFLNVLGSMCORLSMONGSYFPGAKGSR-----LCPEG 222

Db 693 KYAEMAGELDKKVAAPAFLOVL-----TDEGHNRCRIILAAALCKIDN 756

QY 223 WFSQCGPDMDSCLSRHSINPYSNRSRLIFSTWINDHIEKKRTIIP-----LYEAIK 277

Db 737 WL-----LKKHAYKIVKSKASKALFYSYH-GHYIQRK---YPTVNLSTLANTLN 781

QY 278 EDGREGVDWEY-FYGLL-----FTSENKLVHIVCHKKTHKLNCD 317

DB 782 SNYAEVNFMLSLGLIGSMHSGVLIIRALTSKNOK---IKQALESLEKNCD 831

RESULT 20

E86527

CT234 hypothetical protein (imported) - Chlamydia pneumoniae (strain J138)

C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C/Date: 02-Mar-2001 #sequence_rev1510 02-Mar-2001 #text_change 23-Mar-2001

C/Accession: E86527

R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise

Nucleic Acids Res. 28, 2311-2314, 2000

A/Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.

A/Reference number: A86491; MUID:20330349; PMID:10871362

A/Accession: E86527

A/Status: preliminary

A/Sequence type: DNA

A/Source: 1925<STO>

A/References: GB:BA000008; NID:98978667; PIDN:BAA98503.1; GSPDB:GN00142

A/Experimental source: strain J138

C/Genetic: CPJ0293

C/Superfamily: Chlamydia trachomatis hypothetical protein CT234

Query Match 5.1%; Score 91.5; DB 2; Length 925;

Best local similarity 21.8%; Pred. No. 18;

Matches 77; Conservative 44; Mismatches 135; Indels 97; Gaps 17;

DB 28 EVLRKGLFQULPERGSRILCY--EDG---TELTEDYFSPVSDNAELVLLT--GQAMQG 80

DB 513 ELTKRWTSTIFPHPAIYASAIHLVFAEHDLHITHIADLDYVGDRLAAIILVRRORAYG 572

DB 81 YVSDI--RRFLSAFHEPQVGLIQAAQQLCEDQAPORORLADLLHNVSONIAET--- 134

DB 573 PYRDLADKRLKELNSDQPEDIVMGLTILKLEKNPQNFPIILDFTNKNEIDILVTCAL 632

DB 135 ---RAEDPPMFEGLIESRQ---SKSGYLRYSC---SRIRSYLREVSYPST-- 177

DB 633 HTSVRANHRKPYCEPLKRLRQCSHNDASQYLKLTISIALDISFVKDLMTTSOLKNTSR 692

DB 178 -----VGAQAQEFRLVIGSMQRLRSMQYNGSYFDRGAKGSR-----LCPEEG 222

DB 693 KYAEAMIGELDKVAPAPLOVL-----TDEGTHNRCRIIAAKALCKIDN 736

DB 223 WFGCGPFDMDSCLSRHSINPYSNRESRIIFSTWINDHIEKKRTIIP-----LVEAIK 277

DB 737 WL-----LKNHAKIVKSKAKALFYSH--GHYIOKK---YPIYNLSLANTLN 781

DB 278 EODGREVDMEY-FYGLL-----FTSENKLVHIVCHCKTKTKLNCND 317

DB 782 SNYAEVNFMLSLGLIGSMHSGVLIIRALTSKNOK---IKQALESLEKNCD 831

Search completed: May 26, 2003, 15:23:48

Job time : 40 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 26, 2003, 14:49:22 ; Search time 65 Seconds
(without alignments)
1071.445 Million cell updates/sec

Title: US-09-748-451-2

Sequence: 1789
1 MLQPKSVKRLALRSPKRCG.....SRVYPTRLRKQPVAKRQ 338

Gapop 10.0 , Gapext 0.5

Search table: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

SPREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1043	58.3	333	13	08UVS5
2	941	52.6	333	13	09DGL7
3	419	23.4	80	4	096P72
4	419	23.4	110	4	096P74
5	419	23.4	115	4	096P73
6	321	17.9	450	5	09V3H0
7	320	17.9	450	5	09V3H0
8	105.5	5.9	571	10	048981
9	104.5	5.8	284	11	09CZL3
10	104	5.7	1540	6	029433
11	102.5	5.7	266	5	09U786
12	98.5	5.5	2629	6	029021
13	98.5	5.5	684	16	082065
14	98.5	5.5	684	16	082787
15	98	5.3	594	11	082238
16	96.5	5.4	271	4	060819

17	96.5	5.4	571	16	08UKH9	08UKH9 agrobacteri
18	95	5.3	383	11	088305	088305 mus musculu
19	94	5.2	1572	11	0885N0	0885N0 mus musculu
20	93.5	5.2	397	8	002689	002689 podopora a
21	93	5.2	481	5	09U787	09U787 drosophila
22	92.5	5.2	266	5	09V652	09V652 drosophila
23	92	5.1	410	4	096M12	096M12 homo sapien
24	91.5	5.1	395	16	067558	067558 aquifex aeo
25	91.5	5.1	458	16	08UVK2	08UVK2 xenopus lae
26	91.5	5.1	925	16	09U516	09U516 chlamydia p
27	91.5	5.1	1022	15	09Z8P5	09Z8P5 chlamydia p
28	91.5	5.1	1022	15	090317	090317 chlamydia p
29	91	5.1	1719	16	09A5H9	09A5H9 caulobacter
30	91	5.1	1719	16	09A5H9	09A5H9 drosophila
31	90.5	5.1	265	11	09U1T3	09U1T3 rattus norv
32	90.5	5.1	340	15	08U0N4	08U0N4 chimpanzee
33	90.5	5.1	986	15	057059	057059 chimpanzee
34	90.5	5.1	1022	15	089620	089620 chimpanzee
35	90.5	5.1	1056	15	004097	004097 arabidopsis
36	90	5.0	1044	10	09Z0M0	09Z0M0 schizosacch
37	90	5.0	1842	3	096WT6	096WT6 human immun
38	89.5	5.0	337	15	09WHZ0	09WHZ0 human immun
39	89.5	5.0	340	15	08U0N3	08U0N3 chimpanzee
40	89.5	5.0	706	5	0950A8	0950A8 caenorhabdi
41	89.5	5.0	1022	15	088135	088135 chimpanzee
42	89.5	5.0	1022	15	087965	087965 chimpanzee
43	89.5	5.0	1148	11	061372	061372 mus musculu
44	89.5	5.0	4116	5	096958	096958 drosophila
45	89.5	5.0	4120	5	09XYD1	09XYD1 drosophila

ALIGNMENTS

RESULT 1	ID	Q8UVS5	PRELIMINARY	PRT	333 AA
AC	Q8UVS5	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	DE	Caspase-activated DNase.			
GN	GN	CAD.			
OS	OS	Gallus gallus (Chicken).			
OC	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;			
OX	OX	NCBI_TaxID=9031.			
RN	RN	[1]			
RP	RP	SEQUENCE FROM N.A.			
RX	RX	Pubmed-11577114.			
RT	RT	Samejima K., Tone S., Earnshaw W.C.			
RT	RT	Cleavage and Stage I Chromatin Condensation for High Molecular Weight DNA			
RT	RT	U. Biol. Chem. 276:45427-45432(2001).			
RL	RL	EMBL; AF0406761; ALU37716.1; ..			
DR	DR	InterPro: IPR003508; CAD.			
DR	DR	Pfam: PF02017; CIDe-N; 1.			
DR	DR	SMART: SM00266; CAD; 1.			
SO	SO	SEQUENCE 333 AA; 38474 MW; 98069CD287755E11 CRC64;			
Query Match		58.3%; Score 1043; DB 13; Length 333;			
Best Local Similarity		60.1%; Pred. NO. 7.8e-86;			
Matches 197; Conservative		51; Mismatches 74; Indels 6; Gaps 2;			
OY	OY	6 KSVIRALRSPKRVAGRSCEVLRKCLRFQDPERGSRILCYEDGTETEDYPSVVD 65			
DB	DB	6 RGFRLRPPGSAQKFGAAGSLRGLRKRCRLQLPLAGSRILCYEGGTETLSAEPRTLP 65			
OY	OY	66 NAEVYLLTLCQAWGQVYSDIRPFLSAFEPQVGLIOAAQQLCDEAPQORRLADLLHN 125			
DB	DB	66 QTEVYLLRPPESWPGCGGVERFLAALCSRIDAVAEARRLRDEBAPRRORLADLLHN 125			

[illegible]

RESULT 2
OBJECT 7

	PRELIMINARY:	PRT:	333 AA.
09DGL7;			
01-MAR-2001 (trEMBLrel. 16, Created)			
01-MAR-2001 (trEMBLrel. 16, Last sequence update)			
01-MAR-2002 (trEMBLrel. 20, Last annotation update)			
Caspase-activated Dnase.			
Brachydanio rerio (zebrafish)			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
Cyprinidae; Danio.			
NCBI_TaxID=7955;			
[1]			
SEQUENCE FROM N.A.			
Liu C.C., Huang C.J.;			
Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.			
[2]			
SEQUENCE FROM N.A.			
Liu C.C., Huang C.J.;			
"Danio rerio caspase-activated Dnase.";			
Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.			
EMBL; AF286179; AAF99706.1; -;			
EMBL; AF426516; AAL40264.1; -;			
HSSP; 000273; 1IBX.			
InterPro; IPR003508; CAD.			
Pfam; PF02017; CIDE-N. 1.			
SMART; SM00266; CAD. 1.			
SEQUENCE 333 AA; 38340 MW; 7E15D76E8824F0FB CRC64;			

Query Match	52.6%	Score 941	DB 13	Length 333
Best Local Similarity	51.5%	Pred. No. 1.2e-76		
Matches 169	Conservative 67	Mismatches 88	Indels 4	Gaps 2

```

QY      64 PDNAELVLLTLGAWOGVSDIRRELSAFHEBPQVLGAAQOLLCDEQAPQORILLADLT
      111 : 1111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dh      66 PDNTDVLTLPHGGSWNGFPADEINRYLGLDRNTEBELLVSAAGLLSDSESPRRRLITGLD
      111 : 1111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      124 HNYSONIAETRAEDPPWFEGLESFRQKSSGLTKRSCSSRIRSYLREVSISPSIV-GAEA
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dh      126 SNLRDSELENIRLODHDHMEGIDITFRKTKSAAMKNCESRIRGYLKEVDGYOTIRKAKT
      11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      183 QEEFLRLVLSMOCRLSMOYNGSYFDRGAKGGSRLCTEGEFGSCOGPPMDSCLSRSRIN
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dh      186 KSEYKKAVALAEKLEAKAARYNGTYIDRSKQANRLCTEGEFGSCOGAFDENSCNFLHSIN
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      243 PYSNRESRLFTSWNLDHIIIEKKRTIPLTVALEIKEDRGDEWDWEYEGLLFTSENKLY
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dh      246 PYGNRESRLFTSWNLDHIIIEKKRPVITPALAKALEANKSNDIINVDFYKLLFTRENKLY
      111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      303 HLYCHKKTTHKLNCDSPRSIRYKQOTLKR 330
      1111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 306 HIVCHKGAHELSCDSRKIY---RVKK 330

RESULT 3

ID	PRELIMINARY;	PRT;	80 AA.
AC	Q96P72		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	DNA fragmentation factor B truncated form III.		
DE	DEP.		

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.
NCBI_TaxID=9606;
OX

RA Bayasasas J.R., Yuste V.J., Perera R., Comella J.X.:
RT "Characterization of splice variants of human caspase-activated DNase"
RT that show CIDE-N structure and function."
RL Submitted (Aug-2001) to the EMBL/GenBank/DBD databases.
DR EMBL; AF409062; AL02007.1;
DR InterPro: IPR003508; CAD.
DR Pfam: PF02017; CIDE-N: 1.
DR
SQ SEQUENCE 80 AA; 9052 MW; 9EB6C9C0B35AE529 CRC64;

```
SQ SEQUENCE 80 AA; 9052 MW; 9EB6C9C0B35AE529 CRC64;
```

Query Match	23.48;	Score 419;	DB 4;	Length 80;
-------------	--------	------------	-------	------------

Matches	80;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

[illegible]

QY	61	PSVPDNAELVLTIGAWQG	80
Db	61	PSVPDNAELVLTIGAWQG	80

RESULT 4

ID	PRELIMINARY;	PRT;	110 AA.
Q96P74;			
AC	Q96P74;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	DNA fragmentation factor B truncated form I.		

OS Homo sapiens (Human).

0C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
0C Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
0X NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.

RT "Characterization of splice variants of human caspase-activated DNase

Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases

DR InterPro; IPR003508; CAD.

SQ SEQUENCE 110 AA; 12106 MW; 34E72888A0F3B909 CRC64;

Query Match	23.48;	Score 419;	DB 4;	Length 110;
-------------	--------	------------	-------	-------------

Matches	80;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

QY 1 MLOKPSVKLRALRSPRKFGVAGRSCQEVLRKGLRFQLPERGSRILCLYEDTETEDYF 600

Db 1 MLQPKSVKLRSPRKFGVAGRSCQEVLRKGCLEQLPERGSRCLCYEDGETELTDYF 600

QY 61 PSVPDNAELVLLTLGQAWQG 80

[illegible]

DE	01-DEC-2001 (TREMblrel. 19, last annotation update)
DB	NBS-LRR type resistance protein (Fragment).
GN	R1.
OS	Oryza sativa (Rice).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC	Ehhardtiales; Oryzaceae; Oryza.
OX	NCBI_TaxID=4530;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=CV_NIPONBARE;
RX	MEDLINE=96081880; PubMed=9419382;
RA	Leister D., Kuth J., Laurie D.A., Yano M., Sasaki T., Devos K.,
RT	"Rapid reorganization of resistance gene homologues in cereal
RL	genomes";
RM	Proc. Natl. Acad. Sci. U.S.A. 95:370-375(1998).
DR	EMBL; AF032688; AAB96985.1; -.
DR	InterPro; IPR000767; Disease_resist.
DR	InterPro; IPR001611; LRR.
DR	InterPro; IPR002182; NB-ARC.
DR	Pfam; PF00560; LRR_3.
DR	Pfam; PF00931; NB-ARC_1.
DR	PRINTS; PRO0364; DISEASESRIST.
FT	NON_TER 1
FT	NON_TER 571 571
SQ	SEQUENCE 571 AA; 64503 MW; 128BA6734F38B2E6 CRC64;
Query Match	5.9%; Score 105.5; DB 10; Length 571;
Best Local Similarity	20.7%; Pred. No. 0.56;
Matches 69; Conservative 38; Mismatches 97; Indels 129; Gaps 13.	
Y	43 GSRLCLYEDGTELEDEDYFPYSV-----PNAELVLTLTGQAAGGVSDIR--FLSA 91
Db	240 GLSLC-----TKDTEDDKWNLRSEIWEPLSDKNNIILPALRLSYHHLPAALKRCFAFCSV 294
Y	92 PHEQYV-----GLQAAGQLCDPE-----QAQRQLADLL 123
Db	295 FHKQYVEKETLVQIMMALGFQSPGRTEIELGSSYFDLIGRSFFHNHGYYMDAM 354
Y	124 HNVSQNTAAE--TRAEPDPWEGLESFRQSRYLSYCESRIISYLEREVSSYSTVGAE 181
Db	355 HDLAQNSVMDCLLDPP-----NSSSTRSRSHLSFECHNRSKR----- 395
Y	182 AQEEFLRYLGSRCORLNSMOYNGSTFDBGANGSRCLTPBGWFSCQGPFDMDCSLSRHSI 241
Db	396 SFEEFL-----GFKKARTLLLNQYKSRSTPSIPSD----- 425
Y	242 NPYSNRESRLFTSWNDLIIEKKRTIPTLVEAIKEODGEVMEYE----- 289
Db	426 -----LFMLKRLHYVELNRRODTTELDPDSI-----GMLKMLRYLNLSGTGYLIPS 471
Y	290 YGLETFSENKLVIHVCHKRTTHKLCNPSPRI 321
Db	472 SIGRLFNLQTLKL-----KNCHVLECIPGSI 497
RESULT 9	
QCZL3	PREDIMINARY; PRT; 284 AA.
ID	QCZL3;
AC	QCZL3;
DT	01-JUN-2001 (TREMBLrel. 17, created)
DT	01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT	01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE	11 days embryyo cDNA, RIKEN full-length enriched library,
DE	clone:2700062L12, full insert sequence.
GN	VIM.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Scuriongnathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Aachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H. A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L. M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M. F.,
 RA Brownstein M. J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D. A., Kaniwa M., Lee N. H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mochizuki N.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K. F.,
 RA Suzuki H., Toyooka K., Wang K. H., Weitz C., Whitaker C., Wilming L.,
 RA Yoshida K., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.,
 RA Functional annotation of a full-length mouse cDNA collection.";
 RL Nucleotide 409,685-690(2001).
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
 DR EMBL: AC012469; BAB28263.1; -.
 DR MCD: MGI:98932; Vim.
 DR InterPro: IPR001664; IF.
 DR InterPro: IPR002957; Keratin_I.
 DR Pfam: PF00038; filament_1.
 DR PRINTS: PRO1248; TYPEKERATIN.
 DR PROSITE: PS00226; IF, 1.
 KW Coiled coil; intermediate filament.
 SQ SEQUENCE 284 AA; 33196 MW; C33A4C3E54FC947E CRC64;
 Query Match 5.8%; Score 104.5; DB 11; Length 284;
 Best Local Similarity 20.2%; Pred. No. 0.27;
 Matches 61; Conservative 44; Mismatches 114; Indels 83; Gaps 12;
 QY 38 OLPERGSRCLYEDGETETEDYPPSPVDAEVLTLTGAMQGYSDIRRLPSAFHPQV 97
 DB 2 RLREKLOEMLQREAEASTQSFQPDVNSLRLDLERKVESIQEEI-ATLKIHDEEI 60
 QY 98 GLTQAA-----QQLCDQAPQORQLADLHVNSONIAERADPPWF----- 142
 DB 61 QELQAOIQEQHVQIDVYKRDPLALRDVRRHY-BVAAKMLQAEEMYSKPADLSEA 119
 QY 143 -----EGLEFRQSGSYLR-----SCESRIR-----SYLRVSSYST 177
 DB 120 ANRNDALQAKQESNEYRRQVOSLTCEVDALGCTHESLEFRQREKEMFALFANTQDT 179
 DB 178 VQAEAOEFLRVLSGMCORLSMOYNGSTFDRGAKGSRCLTPGWSGCGPDMDSCLS 237
 DB 180 IG-RLDDETQNKKEEMARHLREYO-----DLNVKMLDLIATY 218
 QY 238 RSHINPYSNRESRI-----LPSTWNLDT-----TEKRTIIPLVLEAKRQDGRE 283
 DB 219 RKLL--EGEESRISLPLPTFSLSLNRKTNLESLPLVDTHSKRTL--LTKVTETRDQV 272
 QY 284 VD 285
 DB 273 IN 274
 RESULT 10
 Q29433 PRELIMINARY; PRT; 1540 AA.
 AC Q29433;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE Apolipoprotein B (Fragment).
 GN APOB.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_Taxid=9823;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89108006; PubMed=2905687;
 RA Meda N., Ebert D.L., Doers T.M., Newman M., Hasler-Rapacz J.,
 RA Attie A.D., Rapacz J., Smithies O.,
 RT "Molecular genetics of the apolipoprotein B gene in pigs in relation
 RT to atherosclerosis.";
 RL Gene 70:213-229(1988).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89108006; PubMed=2905687;
 RA Meda N., Ebert D.L., Doers T.M., Newman M., Hasler-Rapacz J.,
 RA Attie A.D., Rapacz J., Smithies O.,
 RT "Molecular genetics of the apolipoprotein B gene in pigs in relation
 RT to atherosclerosis.";
 RL Gene 69:213-229(1988).
 DR EMBL: M20384; AAA64300.1;
 DR EMBL: M22647; AAA30997.1; -.
 KW Lipoprotein.
 KW NON_TER
 SQ SEQUENCE 1540 AA; 174656 MW; 1B67510A18582F0E CRC64;
 Query Match 5.8%; Score 104; DB 6; Length 1540;
 Best Local Similarity 19.5%; Pred. No. 2.9; 123; Indels 98; Gaps 14;
 Matches 66; Conservative 52; Mismatches 123;
 QY 36 RQQLPER-----GSRCL-----LY-----EDGETETEDYPPSPVDAEVL----- 69
 DB 1166 RQQLPERARNTYGDLDLQVMTVEGEVLSQYISKHSGLEILLSTFQDMERKSLNKA 1225
 QY 70 -----VLLTLGAMQGYSDIRRLPSAFHPQVGLTQAQQLC----- 108
 DB 1226 IKFTFDSYKQLTDVYSEYGEOLKLSQDVQKLSLHINTTSLTSLQFLGIRFQET 1285
 QY 109 -DEAPQORQLADLHVNSONIAERADPPWFEGLESPRQSGSYLRSCSRI-- 164
 DB 1286 EEEELRLKREKRLADINDIIONINTYAPLGFRLKENIDSPFGMLNEFIQNTLMEA 1345
 QY 165 RSYLRVSSY-----PSYVG-----AEAOEFLRVLSGMCORLSMOYNGSYFDR 209
 DB 1446 SQELQOLHGYIALKREYFDSMVGWYKYELEERVIMLNLDVYK--DFSKRYTVS 1403
 QY 210 GAKGSRCLTPGWSGCGPDMDSCLS-----RSHINPYSNRESRIIPSTWNLDT 260
 DB 1404 ATQFASQLSSQVEQVQ-----DIQELVSLIADADGKREKIALSLRAQEIIT-KSMV-- 1457
 QY 261 IIEKRTIIPLVLEAKRQDGREVDWEYFGLLTSEN 299
 DB 1458 -----TKKEIISNTHQR-----FYKILDFDQ 1481
 RESULT 11
 Q90786 PRELIMINARY; PRT; 266 AA.
 ID Q90786;
 AC Q90786;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE DREP3.
 GN DREP3 OR DREP3 OR CG8364.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 GN NCBI_Taxid=7227;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20085845; PubMed=10627165;
 RA Inohara N., Nunez G.,
 RT "Genes with homology to DFF/CID5 found in drosophila melanogaster.";
 RL Cell Death Differ. 6:823-824(1999).

DR EMBL: AF149796; AAF03219.1; -
DR HSP: Q9UHD4; 1048.
DR FLYbase: FBgn0028407; Rep3.
DR InterPro: IPR003508; CAD.
DR Pfam: PF02017; CIDE-N; 1.
DR SMART: SMO0266; CAD; 1.
SO SEQUENCE 266 AA; 29414 MW; 5A9757BCA8AFA645 CRC64;
Query Match 5.7%; Score 102.5; DB 5; Length 266;
Best Local Similarity 25.5%; Pred. No. 0.37;
Matches 42; Conservative 30; Mismatches 72; Indels 21; Gaps 6;
QY 2 LQKSKVKRLALRPRKFGVAGRCQVLRKGCRLQPLPERSLCLTEDTEL-TEDYF 60
DB 115 LQNSKPRKIDITNIRKAVATLSELRTKVSLSKFERAQR-RLHLDCDGETVDDEEYF 172
QY 61 PSVPDNEELVLLTGAQMO--GYVSDIR-FLSAFHEPQVGLIQAAQQLDCBOAPQ 116
DB 173 STLEPNMELIAVEGEQWRDSDYNANIRKTSLSA-----QRLSLVSKLQPNYMN 223
QY 117 RLADLLHNVSQNTAETRAEDPWFEGLESFQSKSGYLRYSC 161
DB 224 DDDDKLSNMDPNLSLVDITGREPK---DNEYSARSDAARLSTE 263
RESULT 12
Q29021 PRELIMINARY; PRT; 2629 AA.
AC Q29021;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Apolipoprotein B (Fragment).
GN APOB.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94014802; PubMed=8409766;
RA Puteill C., Maeda N., Ebert D.L., Kaiser M., Lund-Katz S.,
RA Sturley S.L., Kodoyianni V., Grunwald K., Nevin D.N., Aiello R.J.,
et al.
RT "Nucleotide sequence encoding the carboxyl-terminal half of
RT apolipoprotein B from spontaneously hypercholesterolemia pigs.";
RL J. Lipid Res. 34:1323-1335(1993).
DR EMBL: U11235; AAA74655.1; -
DR InterPro: IPR000510; Oxidized_nitrogsel.
DR Pfam: PF00148; oxidized_nitro; 1.
KW Lipoprotein.
NON_TER 1
SEQUENCE 2629 AA; 300092 MW; 5389C4CC17A87049 CRC64;
Query Match 5.6%; Score 99.5; DB 6; Length 2629;
Best Local Similarity 18.7%; Pred. No. 15;
Matches 72; Conservative 61; Mismatches 140; Indels 113; Gaps 16;
QY 36 RFQLEPER-----GSRLC-----LY-----EDGETELTEDYFSPVNAEL----- 69
DB 225 RFQLEPGARNVYTGDELGNMWTVEGVLSQYISKIHSGLEILLSTYFDLAKRKSNTAKR 2314
QY 70 -----VLLTGAQMOGYSDIRRLISAFHEPQVGLIQAAQQLC----- 108
DB 2315 IKFTFDSVKYQTDVYSEYEQSLSLSDQYOKALSDHSINITIELSELQIFLEGIFOEI 2374
QY 109 -DEOAPORORLADLLHNVSQNTAETRAEDPWFEGLESFQSKSGYLRYSCSRI--- 164
DB 2375 EEEETRLKEKRLADFDINDIQONTITNTIYAPLGEFLIKENLDSPFGLNFEIQTLMWEA 2434
QY 165 RSTLYREVSY-----PSTVG-----AEOEELRYLQSGCORLSRMOYNGSYEDR 209
DB 2435 SQELQQLHQYIKALRKREYFPDPSMVGWVKKYVELEEKYINLIKNLVDYVK--DFHSKTYVS 2492

QY 210 GAKGSRCLTPREGWFCGPPMDSCIS-----RHSINPYSNRESRLFTSWNLDH 260
DB 2493 ATDFASQSSQVEQFVE---DIOEYLSILADADGKREKIAQSSRAOEII-KSMAY-- 2546
QY 261 IIEKKRTIIPLYVEI-----KRDGGEVMEWERYGGLTFSENKLVHYCHK--- 308
DB 2547 -----TWKEIISDYHOQFIYKLDPSDQSLDYVE--KFLAETERLIDLSIOXYHM 2594
QY 309 -----KTHKLNDPSRIYKPPQRL 328
DB 2595 FLRYIMKLKLELOSDYVDMRPYIKV 2620
RESULT 13
Q82065 PRELIMINARY; PRT; 684 AA.
AC Q82065;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE DNA helicase IV (EC 3.6.1.-).
GN HELD OR STM1075.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmomella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=172 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Portwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
DR EMBL: AE008746; AAL20008.1; -
DR InterPro: IPR000212; UYRD-helicase.
DR Pfam: PF00580; UYRD-helicase; 1.
KW Helicase; Hydrolase; Complete proteome.
SO SEQUENCE 684 AA; 78072 MW; A77699E4FC2989F CRC64;
Query Match 5.5%; Score 98.5; DB 16; Length 684;
Best Local Similarity 22.9%; Pred. No. 3.1;
Matches 89; Conservative 52; Mismatches 148; Indels 99; Gaps 22;
QY 26 CQEVLRKGCRLQPLPERSRL---CLYEDG-TETLEDYFSPVND-----NAEL 69
DB 152 CREIYMRK-CLAWLQDSEGRQOHQAYADAMLEAHADFFQIIESSPLNPSQARAVNGES 210
QY 70 VLLTGAQMOGYSDI---RRLISAFHEPQVGLI-----QAQQLDCBOAPQ---Q 116
DB 211 SLVLVLAGSGSKTVLVARAGWLLRAGQADGQITLLAFGRKAEEM--DERIERLRLTE 268
QY 117 RLADLLHNVSQNTAETRAEDPWFEGLES-----FQSKSGYLRYSCESR---IRS 166
DB 269 EITARTFHSIALYI-IDGSKKAPVSKLESDARARQDLPLRTROOCSEKKAQAKWRO 327
QY 167 YLREVSYSTVGAEADBEFL--RVLQSGCORLSRMOYNGSYEDRGAAGSGSLCTPGWF 224
DB 328 WLEEMQWVVEGFWDETQRLRLAELRWVSLMHG-----GAQAEKINAGAPE--- 379
QY 225 SCQGPMDSCLSHSINPSN---RESRLFTSWNLDH-----IIEKKRTIIP----- 270
DB 380 ECRLEFGRIKILMAPLAKMSALKAEANAFES--GLHQAQVLLERGRFSPKHLIYD 437
QY 271 -----TLVAKIEODGREV-----DMEFYGLFTSENKLVHYI----- 305
DB 438 EFDIISPORALLALAKRQNSQTTLFVAGDQMAIY--RFGAQDLSLTAFHOFPGGEH 495

```

OY 306 CHKTTTKINDPSRIYKPTRLKRP 333
DB 496 CHDPTTYRNN--SRIGDIANRFVOQNP 520

RESULT 14
ID 0827R7 PRELIMINARY; PRT; 684 AA.
AC 0827R7;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Helicase IV (75 kDa helicase).
GN SRY1097.
Salmonella typhi.
Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
NCBI_TaxID=601;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RA MEDLINE=21354947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moulie S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18."
RL Nature 413:848-852(2001).
DR EMBL; AL627269; CAD08201.1;
DR InterPro; IPR000212; UvrD-helicase.
DR Pfam; PF00580; UvrD-helicase; 1.
KW Helicase; Complete proteome.
SQ SEQUENCE 684 AA; 78038 MW; 5E9077CD302A198E CRC64;

Query Match 5.5%; Score 98.5; DB 16; Length 684;
Best Local Similarity 22.9%; Pred. No 3.1; Mismatches 148; Indels 99; Gaps 22;
Matches 89; Conservative 52;

OY 26 COBYLRKGCIRPOLPERGSL--CLYEDG-TELTEDYPPSPVD-----NAEL 69
DB 152 CREIWER-CLAMLODESSESRQOHQAYADMLAHADFTQIESPLNPQARVYNGES 210
70 VLLTLAGAQCYSVDI--RRFLSAFHEPQVGLI-----QAQOQLCDEQAPOR--Q 116
211 SILVLAGAGSGKTSVLAACAGMLARQADAGQILLARGRKAAEEM--DSRIRRLHTE 268
117 RLADLHNVSQNTAETRAEDPPWEGLES-----RFQSKSGYLRYSCESR-----IRS 166
269 ETTARFHFSLALYI-10QSGSKAPVYSKLESDATARHQLFLHTWRQCSSEKKAQAKGRQ 327
167 YLRVSSYPTVGAQAEFL--RVLSMCQRLSRMQVSGSYFDGAGGSRCTPREGWF 224
328 MLEEKQWVYVPEGNWDETLORRLARLDRLRWYSLRMHG-----GAQAEIAGAPE-- 379
OY 225 SCQGFDMDSCLSRHSINPYSN--RESRLTFTFWLUD--IIKKRTIIP----- 270
DB 380 EGRLEFGRIKIMAVLAKWSAIAKNADPE--GLIHQAVILLEKGRFLSPKHLIYD 437
271 -----TLVEYIKRQDGRV-----DMREYFGLLFTSENLKLYHIV----- 305
DB 438 EFQDISPQRAALILKRNQNSQTLTFLVAGDQWQALY--RFSGAQLSTTAAHQTFEGEGE 495
OY 306 CHKTTTKINDPSRIYKPTRLKRP 333
DB 496 CHDPTTYRNN--SRIGDIANRFVOQNP 520

RESULT 15
OY 062238 PRELIMINARY; PRT; 594 AA.
DB 062238;
AC 062238;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Sex-limited protein SLP(w7) alpha-gamma chain (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
GN NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.D2(C4)(H)SLP(A);
RA MEDLINE=86176748; PubMed=3008092;
RA Hemmaway C., Kalit M., Stavenhagen J., Walthall D., Robins D.;
RT "Sequence comparison of alleles of the fourth component of complement
RT (C4) and sex-limited protein (SLP)."
RL Nucleic Acids Res. 14:2539-2554(1986).
DR EMBL; X06454; CAA29760.1;
DR InterPro; IPR001599; MacrogloblnA2.
DR Pfam; PF00207; AZM; 1.
DR Pfam; PF01759; NTR; 1.
FT NON_TER 1 1
FT CHAIN 304 303 POTENTIAL.
FT CHAIN 304 594 POTENTIAL.
SQ SEQUENCE 594 AA; 66454 MW; 9465C569DA37768 CRC64;

Query Match 5.5%; Score 98; DB 11; Length 594;
Best Local Similarity 24.1%; Pred. No 2.8;
Matches 58; Conservative 29; Mismatches 82; Indels 72; Gaps 12;

OY 18 KFGVAGSGCOEV-LRGG--CLRFOLPERGSLICY-----EGTELTEDYPPSPDNNEL 69
DB 327 KLGISGMAIADITLISGFALRGDEKILTSIDRYSHFPTDQPHYL-YFDSVYTRREC 385
OY 70 VLLTLAGAQCYSVDIREFLSAFHEPQVGLQAQOQLCD-----EQAPQORL 118
DB 386 V-----GF-----GASQEVYGLVQSSNAVLDIYSPDHKCVFAATTKSQL 428
OY 119 LADLHNVSQNTAETRAEDPPWEGLESRFQSKSGY-LRYS----- 160
DB 429 LATV--CGSDVCCQACQKCPPLRSLERRVEDKGRMRFACYPYREYGTVAKLRED 485
OY 161 -----ESRISYLRVSSYPTVGAQAEFLRVYLSGMCQRLSRMQVSGSYFDGAG 212
DB 486 GRAAFRLFSKITQVYLHFTDTMASIG-----QTRNFLSRTSCLR-LFPNKELYLMGMD 539
OY 213 G 213
DB 540 G 540

RESULT 16
OY 060819 PRELIMINARY; PRT; 271 AA.
DB 060819;
AC 060819;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D1934617.1.2 (Chloride channel protein CLC-6b) (Isoform 2)
DE (Fragment).
GN CLC6b.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
GN NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=H.;
RA Eritington H.;
RT submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021155; CAA15952.1;

```

[illegible]

DR	EMBL: AE008936; AA145813.1; -	DR	EMBL: AE007884; AA190497.1; -	KW	Plasmid: Complete proteome	SW	SEQUENCE 571 AA; 62030 MW; 9294894EAA5934F7 CRC64;
Query Match	5.4%; Score 96.5; DB 16; Length 571;						
Best Local Similarity	21.6%; Pred. NO. 3.6;						
Matches	79; Conservative 49; Mismatches 122; Indels 115; Gaps 20;						
QY	10 LRAIRSPKFEVAGRSCEVLRKGLRQLEPERGSRLLYEDGTELTEDYPPSYDAAEL 69						
DB	259 LSLVGPSTL--GKHTVEIVREADLVLLGTRFN---NGTDMWRQSP---DAOI 306						
QY	70 VLLTIGAGMGVSDIRRFSLFAFHEPOVGLIQAAQQLCDQAPQORLL-LADL-LHWVS 127						
DB	307 IHIDTDPQEVARNREARILVGDARTLAGL-----REAIRLCDLHKRYS 351						
QY	128 ONIAETRAEDPPMEEGLESFPQSGYLKSCS-----RINSYLRVSSYSTVGA 181						
DB	352 REAVCNRLRES--WKFENDDR---RGY--YSSEASPLRPERIMAELOGVIDENTYVAD 403						
QY	182 AQEELRLVLSM---CORLRSMQYNGSFDDGAG-----GSRLLCTPE----- 221						
DB	404 ASYSMMWLGLRLISSERTVLT-----PRLAGLGVGVLATGAKVAAPSDVALVG 457						
QY	222 -GWSCGPGFMDSCLSRHSINPYNSRSLIFSTW-MLDHRIEKKRIIPLYEAKR-- 277						
DB	458 DGGFA-----HS-----MALEETLY--RMQIPITIVLANG 486						
QY	278 ----EODGREVDWEFTYGLTTSERLKLVLHYCHKTTTKLNCDSRIYKQOTRLKRQ 333						
DB	487 ILGFQDAETVYFKRYTTACHFAE-----VDHYKLIHAAGCAVAVTQDILKAHMK 539						
QY	334 VKRKO 338						
DB	540 VKRRK 544						
RESULT 18							
088305	PRELIMINARY; PRT; 383 AA.						
AC	088305;						
DT	01-NOV-1998 (TREMBLrel. 08. Created)						
DT	01-NOV-1998 (TREMBLrel. 08. Last sequence update)						
DT	01-MAR-2002 (TREMBLrel. 20. Last annotation update)						
DE	Sex-limited protein (Fragment).						
GN	SLP.						
OS	Mus musculus (Mouse).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
RN	NCBI_TaxID=10090;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=F1 ANIMALS OF C57BL6 AND CBA CROSS;						
RA	Matsunoto K., Ikuta T.;						
RT	"Mus musculus 5' truncated pseudogene of tenascin-X, steroid 21-						
RT	hydroxylase (Cyp21), and sex-limited protein (Slp) genes, partial cds						
RL	and complete sequences."						
RL	Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AB015623; BAA31154.1; -						
DR	InterPro; IPR001599; Macroloblin2.						
DR	InterPro; IPR001134; Netrin_C.						
DR	Pfam; PF00207; AZM; 1.						
DR	Pfam; PF01759; NTR; 1.						
FT	NON_TER 1						
FT	SEQUENCE 383 AA; 43649 MW; EE2AA97B15597047 CRC64;						
Query Match	5.3%; Score 95; DB 11; Length 383;						
Best Local Similarity	27.6%; Pred. NO. 2.9;						
Matches	50; Conservative 21; Mismatches 62; Indels 48; Gaps 11;						
QY	18 KFGVAGSCQEV-LRKG--CLRFQPLPERGSRILCY----EDGTELTEDYPPSYDAAEL 69						

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 15:23:03 ; Search time 297 Seconds
(without alignments)
2562,880 Million cell updates/sec

Title: US-09-748-451-2

Perfect score: 1.789
Sequence: 1 MLOKPKSVKLRLALSPKRGK.....SRIVKPTRLKRPVKKRQ 338

ng table:

BIOSUM62	
Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Lasting first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h
-O=/gen2_1/USPTO.spool/US09748451.unal.21052003.153829.16273/app_query.fasta.1.519
-DB=N Geneseq.101002 -OFMT=faststep -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.csl
-LIST=45 -DOCALLIGN=200 -THR.SCORE=DOT -THR.MAX=100 -THR.MIN=0 -ALIGN=20
-MODE=LOCAL -OUTFMT=ptc -NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09748451.ecgn.1.1.263.funat.21052003.153829.16273 -NCPU=6 -ICPU=3
-NO.MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSEBLOCK=100 -LOOPEL=0
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

1: N Geneseq.101002.1
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	1789	100.0	AAH74642	Nucleotide sequenc
2	1789	100.0	AAH02498	Human DNA fragment
3	1786	99.8	AAH19907	Caspase activated
4	1763	98.5	AAH38992	Human DNA fragment
5	1416	79.2	AAZ21173	Mouse caspase-acti
6	1416	79.2	AAZ21180	Mouse caspase acti
7	372	20.8	AA189481	human poly(nucleo
8	321	17.9	AB115179	Drosophila melanog
9	303.5	17.0	AA216100	Human gene express
10	295	16.5	AB115178	Drosophila melanog
11	124.5	7.0	ABR44763	Human CDNA differe
12	106.5	6.0	AAE21607	Neisseria meningit
13	106.5	6.0	AAE21607	N. meningitidis B
14	105	5.9	AAAB1540	Human Vimentin cod
15	102.5	5.7	AAAO9307	Human DNA sequence
16	102.5	5.7	AAAO9307	Porcine acylglucos
17	101.5	5.7	AAAO9307	Mouse ischaemic co
18	101	5.6	ABN51606	Drosophila melanog
19	101	5.6	AB114002	Human polynucleoti
20	98	5.5	AB114002	CDNA #149 encoding
21	97.9	5.4	ABR44763	Human glutamate re
22	97.5	5.4	ABR44763	Human glutamate re
23	97.5	5.4	ABR44763	Human glutamate re
24	96	5.4	ABR44763	Human glutamate re
25	96	5.4	ABR44763	Human glutamate re
26	96	5.4	ABR44763	Human glutamate re
27	96	5.4	ABR44763	Human glutamate re
28	95.5	5.3	ABO73089	Human osteoblast d
29	95.5	5.3	ABO73089	Human osteoblast d
30	95.5	5.3	ABO73089	Human osteoblast d
31	95.5	5.3	ABO73089	Human osteoblast d
32	94.5	5.3	ABO73089	Human osteoblast d
33	94.5	5.3	ABO73089	Human osteoblast d
34	94	5.3	ABO73089	Human osteoblast d
35	93.5	5.2	AAK51968	Human polynucleoti
36	93	5.2	AAK51968	Human polynucleoti
37	92.5	5.2	AAK51968	Human polynucleoti
38	92.5	5.2	AAK51968	Human polynucleoti
39	92.5	5.2	AAK51968	Human polynucleoti
40	92.5	5.2	AAK51968	Human polynucleoti
41	92.5	5.2	AAK51968	Human polynucleoti
42	92.5	5.2	AAK51968	Human polynucleoti
43	92.5	5.2	AAK51968	Human polynucleoti
44	92.5	5.2	AAK51968	Human polynucleoti
45	92.5	5.2	AAK51968	Human polynucleoti

ALIGNMENTS

RESULT 1	
AAH74642	AAH74642 standard; cDNA: 2839 BP.
ID	AAH74642
AC	AAH74642
XX	15-OCT-2001 (first entry)
DT	
XX	Nucleotide sequence of human DNA fragmentation factor 40 (DFF40).
DE	
XX	Human; DNA fragmentation factor; DFF40; DFF45; apoptosis; DNase;
KW	molecular chaperone; cancer cell; ss.
XX	
OS	Homo sapiens.
XX	
XX	
FT	Location/Qualifiers
FT	132..1148
FT	CD5
FT	/*tag - a

/product= "DNA fragmentation factor 40 (DFP40)."

FT XX US2001011078-A1.
 PN XX 02-AUG-2001.
 PD XX 22-DEC-2000; 2000US-0748451.
 PE XX 16-APR-1998; 98US-0061702.
 PR XX (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Wang X, Liu X;
 PI
 XX WPI: 2001-496169/54.
 DR P-PSDB; AAG63592.
 XX
 PT New DNA fragmentation factor polypeptides and polynucleotides, useful
 PT for inhibiting the growth of cancer cells, as well as for inducing
 PT apoptosis of cells -

Claim 18; Page 44-45; 56pp; English.

The present sequence encodes a human DNA fragmentation factor subunit of 40 kDa, designated DFP40. The specification also describes DFP45. DFP40 is capable of inducing apoptosis, and may contain a nuclear localization fragment. DFP45 acts as a molecular chaperone to direct the folding of DFP40. Although all DFP40 activity is associated with DFP40, DFP45 activity only occurs once DFP40 is complexed with DFP45. CC The DFP polypeptides and polynucleotides are useful for inhibiting the growth of cancer cells, and for inducing apoptosis of cells.

XX Sequence 2839 BP; 644 A; 701 C; 755 G; 739 T; 0 other;

Alignment Scores:

Pred. No.: 1.78e-176 Length: 2839
 Score: 1789.00 Matches: 338
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0

US-09-748-451-2 (1-338) x AAH74642 (1-2839)

QY 1 MetLeuGlnIysPProLysSerValLysLeuArgAlaLeuArgSerProAlaGlyPheGly 20
 DB 132 ATGCTCCAGAAAGCCCAAGAGCGCTGACCTGGCGCCAGCCGAGGAAAGTGGCG 191
 QY 21 ValAlaGlyArgSerGlyGlnValLeuArgGlyGlyGlyGlyGlyGlyGlyGlyGly 40
 DB 192 GTGGCTGGCCGAGAGCTGCGAGAGGCTGCGCAAGGGCTGTCTCCCTCCAGCTCCCT 251
 QY 41 GlnArgGlySerArgLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 60
 DB 252 GAGGCGGGTTCGGGCTGTGCTGTACGAGATGGACGAGCGAGCGAAGATTACTTC 311
 QY 61 ProSerValProAspAsnAlaGlnLeuValLeuLeuThrLeuGlyGlnAlaTrpGlnGly 80
 DB 312 CCCAGTCTCCGCAACAGCCGAGCTGTGCTGCTCACTTGAGGCGCAAGGCTGCAAGGC 371
 QY 81 TyrValSerAspIleArgArgPheLeuSerAlaPheHisGlyProGlnValAlaTrpGlnGly 100
 DB 372 TATGTGAGCGACATCAGCGCTTCTCTCAATTCATTCACGACCAAGGTGGGCTCATC 431
 QY 101 GlnAlaAlaGlnGlnLeuLeuGlyAspGlnGlnAlaProGlnAlaArgGlnAlaLeuAla 120
 DB 432 CAGGCGCCGCCAGCGCTGTGTGTGATGAGAGGCCGCCACAGAGCGAGAGCTGTGCT 491
 QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGlnTrpArgAlaGluAspProPro 140
 DB 492 GACCTCTCGACACAGCTACGACCAACATCGCGCGAGACCCGGCGTGAAGACCCGCG 551
 QY 141 TrpPheGlnGlyLeuGlnSerArgPheGlnSerIysSerGlyTrpLeuArgTyrSerCys 160

DB 552 TGGTTGAAGCTTGGAGTCCCGATTCAGAGCAAGCTGCGCTATCTGAGATACAGCTGT 611
 QY 161 GluSerArgIleArgSerTyrLeuArgGlnValSerSerTyrProSerThrValGlyAla 180
 DB 612 GAGAGCCGGATCCGAGATTACCTGAGGAGGTGAGTCTTACCCCTCCACAGTGGGTGCG 671
 QY 181 GlnAlaGlnGlnGlnLeuValLeuGlySerMetCysGlnArgLeuArgSerMet 200
 DB 672 GAGCTCAGAGAGATTCTCGGGCTCTCGCTCCATGTCCAGAGAGCTCCGGTCCATG 731
 QY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaLysGlySerArgLeuGlyCysThrPro 220
 DB 732 CAGTACATGGCAGCTTCTGACAGAGAGAGCCCAAGGGGCGACCCGCTCTGCACACG 791
 QY 221 GlnGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 DB 792 GAGGCTGTGTTCTCTGCGAGGTTCCCTTTACATGAGACGCTCTTATCAAGACACTCC 851
 QY 241 IleAsnProTyrSerAsnArgGluSerArgIleLeuPheSerThrTrpAsnLeuAspHis 260
 DB 852 ATCAACCCCTACAGTACAGAGAGAGAGAGAGATCTTCAACACCTGGAACCTGGATCAC 911
 QY 261 IleIleGlnLysLysArgThrIleIleProThrLeuValGlnAlaIleLysGlnGlnAsp 280
 DB 912 ATAAATGAGAAAGAAACGACCATCATCTTCACTGGGAGGCAATTAAAGAAACAAGAT 971
 QY 281 GlnArgGlnValAspTrpGlnTyrPheTyrGlyLeuLeuPheThrSerGlnAsnLeuLys 300
 DB 972 GGAAGAGAAAGGAGCTGGAGATATTATTATGGCTGCTTTTACTCAGAGAACCTTAAA 1031
 QY 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSerArg 320
 DB 1032 CTAGTGCACATTTGCTGCGCATTAAGAAACCAACCAAGCTCACTGAGACCCGAGCAGA 1091
 QY 321 IleTyrLysProGlnThrArgLeuArgLysGlnProValArgLysArgGln 338
 DB 1092 ATCTACAAACCCAGACAAAGGTGAAGCGGAGAGCGCTGTGCGAAAGCCGAC 1145
 RESULT 2
 AAD02498 standard; cDNA; 2839 BP.
 AC AAD02498;
 XX 24-APR-2001 (first entry)
 DE Human DNA fragmentation factor 40 (DFP40) cDNA.
 KW Human; DNA fragmentation factor; DFP; apoptosis; molecular chaperone;
 KW gene therapy; hyperproliferative disorder; therapy; tumour; restenosis;
 KW psoriasis; angiogenesis; cancer; cytosolic; neoplasia; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 132..1148
 FT /*tag= a
 FT /product= "DNA fragmentation factor 40 (DFP40)."
 XX
 EN US6165737-A.
 XX
 PD 26-DEC-2000.
 XX
 PF 16-APR-1998; 98US-0061702.
 XX
 PR 16-APR-1998; 98US-0061702.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Wang X, Liu X;
 XX
 DR WPI: 2001-090481/10.

DR P-PSDB: AAY72467.
 XX Identifying modulator of human DNA fragmentation factor 40, for
 PT treating cancer, involves contacting cell or cell-free composition
 PT comprising DFF40 with candidate substance and comparing apoptosis with
 PT control
 XX Example 1: Column 77-80; 52pp; English.
 PS
 CC The present sequence is a human DNA fragmentation factor 40 (DFF40)
 CC CDNA capable of inducing apoptosis. DFF40 acts as a molecular chaperone
 CC comprising 40kDa and 45kDa subunits. DFF40 acts as an inhibitor
 CC to facilitate the appropriate folding of DFF40 and acts as an inhibitor
 CC for DFF40. DFF40 and DFF40 are used in gene therapy. The modulators of
 CC human DFF40 activity are useful for inducing apoptosis and for treating
 CC tumours, proliferative disorders such as metastasis, psoriasis, metastatic
 CC oligodendroglioma and benign and malignant neoplasias. They are also
 CC used for treating cancers of the brain (glioblastoma, astrocytoma,
 CC node, pancreas, small intestine, blood cells, colon, stomach, breast,
 CC endometrium, prostate, testicle, ovary, skin, head and neck, oesophagus,
 CC bone marrow, blood, other tissue and multi-drug resistant cancer.
 CC
 SQ Sequence 2839 BP; 644 A; 701 C; 755 G; 739 T; 0 other:
 Alignment Scores:
 Pred. No.: 1.78e-176 Length: 2839
 Score: 1789.00 Matches: 338
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 22 Gaps: 0
 US-09-748-451-2 (1-338) x AAD02498 (1-2839)
 OY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
 DB 132 ATGCTCCAAAGCCCAAGCGTGAAGCTCGGGCCCTCGAGCCGAGAGAGTTGGC 191
 OY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 DB 192 GTGGCTGGCGGAGTGGCCAGAGAGTCTCGCAAGAGGCTGCTCCCTCAGCTCCCT 251
 OY 41 GluArgLysArgLysLeuTyrGluAspGlyThrGlnLeuThrGlnAspTyrPhe 60
 DB 252 GAGCGCGGTCCCGCTGCTGTACAGAGATGGACAGAGCTGACGAGAGATTACTTC 311
 OY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaTrpGlnGly 80
 DB 312 CCCAGTGTCCCGCAAGACGCGAGCTGTCTCTCAGCTTGGCCAGGCTCGAGAGGC 371
 OY 81 TyrValSerAspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
 DB 372 TATGTGAGGAGATCAGAGCTCTCTCAGTGCATTTCAGAGCCACAGGTGGGCTCATC 431
 OY 101 GlnAlaAlaGlnGluLeuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuAla 120
 DB 432 CAGGCGCCCGCAGAGCTGTGTGTGATGAGAGAGGCCCCACAGAGCAGAGCTGTGCT 491
 OY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaGluThrArgAlaGluAspProPro 140
 DB 492 GACCTCTGCAAGAGTACAGAGACATCGCGCCGAGACCCGGCTGAGGACCCGCGG 551
 OY 141 TrpPheGlnGlyLeuGlnSerArgPheGlnSerArgLysSerGlyTyrLeuArgTyrSerGys 160
 DB 552 TGCTTTGAGGCTTGAGTCCCGATTTCAGAGCAAGCTGTGCAATCTGAGATACAGCT 611
 OY 161 GluSerArgIleArgSerTyrLeuArgLysValSerSerTyrProSerThrValGlyAla 180
 DB 612 GAGAGCCGATCCGAGGTACCTGAGGAGGTGAGTCCCTACACCCCTCCACAGAGGAGG 671
 OY 181 GluAlaGlnGluThrPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200

DB 672 GAGGCTCAGAGAGATTCCTCGCGGTCTCTCCATGTCACAGAGGCTCCGCTCATG 731
 OY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaLysGlyGlySerArgLeuCysThrPro 220
 DB 732 CAGTCAATGTGACACTCTTCGACAGAGAGGCGGAGCCGCTCTCTCAGACACCG 791
 OY 221 GluGlyTrpPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 DB 792 GAGGCTGTGTTCTCTCCCGCAGGCTCTTGTGACATGAGACACTCTTATACACACTCC 851
 OY 241 IleAsnProTyrSerAsnArgLysSerArgIleLeuPheSerThrTyrPheAsnLeuAspHis 260
 DB 852 ATCAACCCCTACAGTAACAGAGAGAGAGAGATTCCTTCACACCTGGAACCTGATCAC 911
 OY 261 IleIleGlyLysLysArgTyrThrIleIleProThrLeuValGluAlaIleLysGluGlnAsp 280
 DB 912 ATTAATGAAAGAAAGAAACGACCATTCATTCCTACACTGTGTGAAGCAATTAGGAACAAGAT 971
 OY 281 GlyArgGlyValAspTrpGlyTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLys 300
 DB 972 GGAAGAGAGAGTGGAGTGGAGTATTTTATGAGCTCTCTTTTACCTCAGAGAACCTAAA 1031
 OY 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnLysAspProSerArg 320
 DB 1032 CTAGTGCACATTTGTGTGCTATAGAAAACACCCACACAGCTGAGTGTGACCCGAGCA 1091
 OY 321 IleTyrLysProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
 DB 1092 ATCTACAAACCCCAAGCAAGGTGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1145
 RESULT 3
 AAX19907
 ID AAX19907 standard; DNA; 1017 BP.
 AC AAX19907;
 XX 11-JUN-1999 (first entry)
 DE Caspase activated nuclease CPAN encoding DNA.
 XX
 KW Caspase activated nuclease; CPAN; apoptosis; human; heart attack;
 KW stroke; tumour; ss.
 OS Homo sapiens.
 XX
 PN W09910501-A1.
 XX
 PD 04-MAR-1999.
 XX
 PF 19-AUG-1998; 98MO-US17214.
 XX
 PR 06-JAN-1998; 98US-0072192.
 PR 22-AUG-1997; 97US-0056904.
 PR 22-AUG-1997; 97US-0056907.
 XX
 PA (CHIR) CHIRON CORP.
 PI
 PI Halenbeck R, Macdonald H;
 DR WPI: 1999-190620/16.
 DR P-PSDB: AAY04123.
 XX
 XX New isolated caspase activated nuclease - used to develop products
 PT which can inhibit or activate apoptosis for use in treating e.g.
 PT heart attack, stroke or tumours
 XX
 PS Claim 1; Page 38; 61pp; English.
 CC The present sequence encodes a caspase activated nuclease (CPAN). Active
 CC CPAN proteins can be used for identifying compounds which inhibit
 CC apoptosis. Inhibitors of apoptosis can be used therapeutically, e.g.
 CC where cells are damaged and the apoptotic pathway is initiated such as
 CC by heart attack or stroke. The inactive CPAN proteins can be used for

CC Identifying compounds which activate apoptosis. Compounds which activate
CC apoptosis can be used to treat tumours. The association of CPAN with a
CC caspase-sensitive inhibitor suggests that when cells enter into
CC apoptosis, the activation of caspases will trigger CPAN to become
CC active.

XX Sequence 1017 BP: 236 A; 293 C; 303 G; 185 T; 0 other:

Alignment Scores:

Prod. No.: 8,34e-177 Length: 1017
Score: 1786.00 Matches: 337
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.70% Mismatches: 0
Query Match: 99.83% Indels: 0
DB: 20 Gaps: 0

US-09-748-451-2 (1-338) x AAX19907 (1-1017)

QY 1 MetLeuGlnIlyProLysSerValIlyLeuArgAlaLeuArgSerProAlqlySphgIy 20
1 ATGCTCCAGAGCCCAAGAGCGTGAAGCTCGGGCCCTGGCGACCCCGAGGAAATTCGGC 60
21 ValAlaGlyArgSerGlyGlnGluValLeuArgGlyGlyCysLeuArgPheGlnLeuPro 40
61 GTGGCTGGCCGGAGCTCCAGAGAGTCTGGCGAGGGCTGTCTCCCTCCAGCTCCCT 120
QY 41 GluArgGlySerArgLeuGlyCysLeuArgIlyAspGlyThrGluLeuThrGluAspPyrPhe 60
121 GAGCGCGGTTCGGCGCTGTCTGTACAGAGATGGACGAGGAGTGCAGGAAGATTACTTTC 180
QY 61 ProSerValProAspAlaGluLeuValLeuLeuThrLeuGlnAlaIlyProGlnIly 80
181 CCAGTGTTCGCCAGCAAGCCGAGCTGTCTGTCTGCTCCTGGCCAGGCTGGCAGGCG 240
QY 81 TyrValSerAspIlyLeuArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
241 TATGTGAGCAATCAGAGCGCTTCTCAGTCAATTACAGACAGCAGAGTGGGCTATC 300
DB 101 GlnAlaAlaGlnGlnLeuLeuGlyAspGlnAlaProGlnArgGlnArgLeuAla 120
301 CAGGCGCCCGCAGCAGCTGTGTGTGATGACAGCCCGCAGAGGAGGCTGTGGCT 360
QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaGluThrArgAlaGluAspProPro 140
361 GACCTCTGCAACAACGTCAGCAGACATCGCGGCGCAGACCCGGGCTGAGAGCCCGC 420
QY 141 TrpPheGlnGlyLeuGlnIlySerArgPheGlnSerIlySerGlyTyrLeuArgTyrSerCys 160
421 TGGTTGGAAGGCTTGGAGTCCCGATTTCAGAGCAAGCTGGCTATCTGAGATACAGCTGT 480
QY 161 GluSerArgIlyLeuArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
481 GAGAGCCGGATCCGGACTTACTGTAGAGGAGGTGAGCTCTTACCCCTCAGCGTGGGTCGC 540
DB 181 GlnAlaGlnGlnIlyPheLeuArgValLeuGlySerMetCysGlnAlaGluLeuAspSerMet 200
541 GAGGCTCAGGAGAAATTCCTGGGGTCTGGGCTCCAGATGTCAGAAAGCTCCGGTCCATG 600
QY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaIlyGlyLeuSerArgLeuGlyCysThrPro 220
601 CAGTACAATGAGCACTCTCGACAGAGAGCAAGGCGGCGGCGCCCTCTGACACCG 660
QY 221 GluGlyTyrPheSerGlyGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
661 GAAGCGGTGCTTCTGCTCCAGGCTCCCTTGTACATGACAGCTCTTATGACACATCC 720
DB 241 IleAsnProTyrSerAsnArgIlySerArgIleLeuPheSerThrTrpAsnLeuAspHis 260
721 ATGACCCCTTACAGTACAGAGAGAGAGATCTTTCAGACACCTGGAACCTGATCAC 780
QY 261 IleIleGlnIlyLysArgThrIleIleProThrLeuValGlnAlaIleIlyGlnIlyAsp 280
781 ATATATGAAAAGAAAGCAGCAGCATCATCTCTTCACTGTGAGAGCAATTAAGAAAGAT 840

QY 281 GlyArgGluValAspTrpGlyIlyPheTyrGlyLeuLeuPheThrSerGlnAsnLeuLys 300
841 GGAAGAGACAGTGGAGCTGGAGTATTTTATGGCTCTGTTTACCTCAGAGAACTTAA 900
DB 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnGlyAspProSerArg 320
901 CTAGTGCACATGTGCTCCATTAAGAAAACCAACCAAGCTCAACTGTGACCGAGCAGA 960
QY 321 IleTyrLysProGlnTrpArgLeuLysArgGlyGlnProValArgLysArgGln 338
961 ATCTACAAACCCAGACAGAGTTGAACGAGAGAGCTGTGGGAAACCCAG 1014
DB 961 ATCTACAAACCCAGACAGAGTTGAACGAGAGAGCTGTGGGAAACCCAG 1014
RESULT 4
AAZ38992
ID AAZ38992 standard; cDNA; 2839 BP.
XX
AC AAZ38992;
XX
DT 22-FEB-2000 (first entry)
XX
DE Human DNA fragmentation factor DFF40 nucleotide sequence.
XX
KW Human: DNA fragmentation factor; DFF40; DFF45; apoptosis; gene therapy;
KW Cytostatic; growth; tumour; ss.
XX
OS Homo sapiens.
XX
FN WC954482-A1.
XX
PD 28-OCT-1999.
XX
PF 16-APR-1998; 98WO-0507895.
XX
PR 16-APR-1998; 98WO-0507895.
XX
PA (TFXA) UNIV TEXAS SYSTEM.
XX
PI Wang X, Liu X;
XX
PI
XX
DR MPI: 2000-052702/04.
DR P-PSDB: AAY57440.
XX
PT DNA fragmentation factor DFF40 involved in apoptosis and related
PT polynucleotide -
XX
PS Claim 18; Page 125-127; 154pp; English.
XX
CC The present sequence encodes a human DNA fragmentation factor,
CC designated DFF40. Also described are: (1) a method of inducing apoptosis
CC in a cell comprising providing the cell with DFF40 which results in
CC apoptosis; (2) a method for inhibiting the growth of a cancer cell
CC comprising contacting a cancer cell with a DNA fragmentation factor
CC designated DFF40 under conditions permitting the uptake of the DNA
CC fragmentation factor by the cell where the presence of the DFF40 into
CC the cell induces apoptosis; (3) a method for treating cancer comprising:
CC (a) encoding a DFF40 DNA fragmentation factor; and (b) a promoter active
CC in the tumour cell, where the promoter is operably linked to the region
CC encoding the DNA fragmentation factor, under conditions permitting the
CC uptake of the nucleic acid by the tumour cell; (4) a method of
CC identifying a modulator of DFF40; and (5) a method of producing a
CC functional DNA fragmentation factor. An expression construct encoding a
CC DNA fragmentation factor DFF40 and DFF45 complex is provided to a cell to
CC induce apoptosis, especially in tumour cells. DFF40 is used to inhibit
CC the growth of a cancer cell, especially in humans.
XX
SQ Sequence 2839 BP: 642 A; 699 C; 750 G; 737 T; 11 other:

Alignment Scores:

Prod. No.: 9,26e-174 Length: 2839
Score: 1763.00 Matches: 334
Percent Similarity: 98.82% Conservative: 0
Best Local Similarity: 98.82% Mismatches: 4

```

Query Match: 98.55% Indels: 0
DB: 21 Gaps: 0
US-09-748-451-2 (1-338) x AA238992 (1-2839)

QY 1 MetLeuGlnIlyProLysSerValIlyLeuArgAlaLeuArgSerProArgIlyPheGly 20
DB 132 ATGCTCCAGAGCCCAAGACGTCGAGGCGCCCTGCGCCAGCCGCGCCANNNNNNNNNN 191
QY 21 ValAlaGlyArgSerCysGlnGluValLeuArgIlyGlyCysLeuArgPheGlnLeuPro 40
DB 192 GTGGCTGGCGGAGCTCCAGAGAGGTGGCGGAGAGGGGTGCTCCGCTCCAGCTCCCT 251
QY 41 GluArgIlySerArgLeuCysIleuArgIlyGluAspGlyThrGluLeuThrGluAspTyrPhe 60
DB 252 GAGCGCGGTCCCGGCTCTGCTCCCTGAGAGAGATGGCAGCGAGCTGACGAGAGATCTTC 311
QY 61 ProSerValProAspAsnAlaGluLeuValIleuLeuThrIleuGlyGlnAlaIleuArgIly 80
DB 312 CCCAGTGTCCCGACACACCGCGAGCTGTGCTGCTCCTTGGCCAGGCGCTGCGAGGCG 371
QY 81 TyrValSerAspIleuArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
DB 372 TATGTGAGCGACATCAGCGCTTCCTCAGTGCATTTTCACAGACCGACAGTGGGCTCATC 431
QY 101 GlnAlaAlaGlnGlnLeuLeuCysAspGlnAlaProGlnIleuArgIleuLeuAla 120
DB 432 CAGGCGCGCCAGCGAGCTGCTGTGTGATGAGCAGGCGCCCGACAGGACAGGCGCTCTGCT 491
QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspPro 140
DB 492 GACCTCTGACACACGTCACAGACGACGAGCGCGGCGGAGCGGCTGAGAGACCGCGCG 551
QY 141 TyrPheGlnIlyLeuGlnSerArgPheGlnSerIlyTyrLeuArgTyrSerCys 160
DB 552 TGGTTGAAGCTTGGAGTCCCGATTTTCAGAGCAAGTGGCTATCTGATGATACAGTGT 611
QY 161 GluSerArgIleuArgSerTyrLeuArgGluValIleuSerTyrProSerThrValGlyAla 180
DB 612 GAGACCGGATCCGAGATCTACCTGAGGAGGTAGTCTCTACCTCCACAGCTGGGCG 671
QY 181 GluAlaGlnGlnIlyLeuLeuArgValIleuGlySerMetCysGlnIleuArgSerMet 200
DB 672 GAGCTCAGAGAGATTCCTGCGGCTCTGCTGCTCATGTGCGCAGAGGCTCTCGGCTCATG 731
QY 201 GlnTyrAsnGlySerTyrPheAspArgIlyAlaIlyGlyGlySerArgLeuCysThrPro 220
DB 732 CAGTCAATAGGACCTACTCTGACAGAGGAGCCAGGCGGCGGCGCTCTGACACCG 791
QY 221 GluGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
DB 792 GAAGCTGTGCTCTCTCCAGAGGCTCCCTTGAACATGACAGCTGTATTCAGACACTCC 851
QY 241 IleAsnProTyrSerAsnArgGluSerArgIleuPheSerThrTyrAsnLeuAspHis 260
DB 852 ATCAACCCCTACAGTACAGGAGAGAGATCTCTTCAGACACTGAGACCTGAGATCAC 911
QY 261 IleIleGluIlyLysArgThrIleIleProThrLeuValGluAlaIleLysGluGlnAsp 280
DB 912 ATATATGAAAGAAAGAACGACCATCATCTCCACACTGTGAGAGCAATTAAGAACAAAGAT 971
QY 281 GluArgGluValIleAspTyrGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLys 300
DB 972 GGAAGAGAAAGTGAAGTGGAGATATTTTATGAGCTGCTTTTACTCTGAGAACCTAAAA 1031
QY 301 LeuValHisIleValCysHisIlyLysThrThrHisIlyLeuAsnCysAspProSerArg 320
DB 1032 CTATGTCACATTTCTGCTCATTAAGAAACCCACCACTCAACTGTGACGCCAGCA 1091
QY 321 IleTyrLysProGlnIleThrArgLeuLysArgIlyGlnProValIleArgLysArgGln 338
DB 1092 ATCTCAAAACCCACAGCAAGGTGATGAGAGACCACTGTGCGGAAAGCGCAG 1145

```

```

RESULT 5
AA221173
ID AA221173 standard; cDNA; 1038 BP.
XX
AC AA221173;
XX
DE 23-NOV-1999 (first entry)
XX
DE Mouse caspase-activating DNase encoding cDNA.
XX
KW Mouse; caspase activating DNase; CAD; Inhibitor; ICAD; ICAD-L; ICAD-S;
XX cell death; ss.
XX
OS Mus sp.
XX
FT CDS
FT 1..1035
FT /tag= a
FT /product= "CAD"
FT /note= "caspase-activating DNase"
XX
XX JP11239495-A.
XX
XX PD 07-SEP-1999
XX
XX 25-DEC-1998; 98BP-0369222.
XX
XX 25-DEC-1997; 97BP-0369356.
XX
XX (OSAB-) 2H OSAKA BIOSCIENCE KENKYUSHO.
XX
XX WPI: 1999-554024/47.
XX
XX P-PSDB; AAY29931.
XX
XX Inhibitor for caspase-activating DNase - useful for preventing cell
XX death and is useful as reagent for cell death
XX
XX Example: Page 15-16; 26pp; Japanese.
XX
XX The present invention describes an inhibitor (ICAD) for caspase-
XX activating DNase (CAD). Two ICAD's are specifically claimed: a long
XX chain form (ICAD-L) and a short chain form (ICAD-S). ICAD and ICAD cDNA
XX can be used in the prevention of cell death, and can be used as a
XX research reagent for cell death. The present sequence encodes mouse
XX CAD which is used in the exemplification of the present invention.
XX
XX Sequence 1038 BP; 236 A; 293 C; 307 G; 202 T; 0 other;
XX
Alignment Scores:
Pred. No.: 3,88e-138 Length: 1038
Score: 1416.00 Matches: 261
Percent Similarity: 88.17% Conservative: 37
Best Local Similarity: 77.22% Mismatches: 38
Query Match: 79.15% Indels: 2
DB: 20 Gaps: 1
US-09-748-451-2 (1-338) x AA221173 (1-1038)

QY 1 MetLeuGlnIlyProLysSerValIlyLeuArgAlaLeuArgSerProArgIlyPheGly 20
DB 10 GTGCTCGGCGACCAACCAATGGTCAATTCGAGCGCTCAAGAGCGCTGCAAGTGGCG 69
QY 21 ValAlaGlyArgSerCysGlnGluValLeuArgIlyGlyCysLeuArgPheGlnLeuPro 40
DB 70 GTGGCTGGCGGAGCTCCAGAGAGGTGGCGGAGAGGGGTGCTCCGCTCCAGCTCCCT 129
QY 41 GluArgIlySerArgLeuCysIleuArgIlyGluAspGlyThrGluLeuThrGluAspTyrPhe 60
DB 130 ATGCGCGGTCCCGGCTCTGCTCCCTGAGAGAGATGGCAGAGGTGACGAGCGACTGCTTC 189
QY 61 ProSerValProAspAsnAlaGluLeuValIleuLeuThrIleuGlyGlnAlaIleuArgIly 80
DB 190 CCGGCGCTTCCCAAGCGAGTGTGATCTCTCATTCCTGACCGGTGGAGAGACTGGCATGGC 249

```

QY 81 TyrValSerAspIleArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
 |||||
 Db 250 TATGTGAGTGAACATCACAGCTTCTCCAGTGTATGATGAGCCACATGCCGGGTCATC 309
 |||||
 QY 101 GlnAlaIleGlnGlnLeuLeuCysAspGlnGlnAlaProGlnAlaArgLeuAla 120
 |||||
 Db 310 CAGGCTGCAGGACGCTGTCTGATGAGCAGGCCCTCCTGAGAAAGCTCTGACC 369
 |||||
 QY 121 AspleuLeuHisAsnValSerGlnAsnIleAlaIleGluThrArgAlaGluAspPro 140
 |||||
 Db 370 GATCTTGTGATCATCGAGCCAGAACATATTAATCTCAAGACCCGGAGCAGACCCATCC 429
 |||||
 QY 141 TrpPheGluGlyLeuGlnSerArgPheGlnSerIleSerGlyTyrLeuArgTyrSerCys 160
 |||||
 Db 430 TGGTTGAAGGTTTGGAGTCGAGATCAGGAATTAAGCGGCTATGAGATACAGCTGT 489
 |||||
 QY 161 GluSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
 |||||
 Db 490 GAGAGTCGATCGAGGCTTACCTTAAGAGAGGTGAGCCCTTACCTTATGCTGATGAA 549
 |||||
 QY 181 GluAlaGlnGluGlnPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
 |||||
 Db 550 GCAGCTCAAGAGAGTACCTGCGAGTCTTGGCTCCATGTGCCAGAACTCAATCGGTG 609
 |||||
 QY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaAlaGlyGlySerArgLeuCysThrPro 220
 |||||
 Db 610 CAGACATAGTCAGAGCTATTTGACAGAGAGTCGAGAGCCAGACGCCCTCTGTACATCA 669
 |||||
 QY 221 GlnGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 |||||
 Db 670 GAAAGAGGTGTTCTCTCCAGAGGCCCTTGTACCTGAGAGCTGTCTTCCAAAGCACTCC 729
 |||||
 QY 241 IleAsnProTyrSerAsnArgGluSerArgIleLeuPheSerThrTyrPasnLeuAspHis 260
 |||||
 Db 730 ATCAACCCCTATGAGCAACAGAGAGAGCCGGATCTCTTCACTGATCTGAACTGATCAT 789
 |||||
 QY 261 IleIleGluGlySerArgThrIleIleProThrLeuValGlnAlaIleGlyGlnAsp 280
 |||||
 Db 790 ATATATAGAGAAGAGCCACCGCTGTGTACCCAGCTGCTGAAAGCCATC-----CAGCAT 843
 |||||
 QY 281 GlyArgGluValAspTrpGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLys 300
 |||||
 Db 844 GGGAGGAGAGTGAAGTGGAGTACTTCTACAGCTGCTCTTCACTGCCGAGAACTCTAAG 903
 |||||
 QY 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSerArg 320
 |||||
 Db 904 CTGGTCACATCGCCTGCACAAAGAGACCAACACAAAGCTGGAGTGCAGACCCGAGTAGG 963
 |||||
 QY 321 IleTyrIleSerProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
 |||||
 Db 964 ATCTATCGCCTCAGACAGATCCAGAGAGAGAGAGAGCTGCTCGAGAGAGCCG 1017
 |||||
 ID AAZ21180 standard; cDNA: 1038 BP.
 AC AAZ21180;
 XX 23-NOV-1999 (first entry)
 DE Mouse caspase activating DNase encoding cDNA.
 KW Mouse; caspase activating DNase; CAD; Inhibitor; ICAD; ICAD-L; ICAD-S;
 KM cell death; cancer; ss.
 OS Mus sp.
 XX Location/Qualifiers
 FH 1..1035
 FT CDS
 FT /tag- a
 FT /product- "CAD"
 FT /note- "caspase activating DNase"

XX JP1239494-A.
 FN 07-SEP-1999.
 PD 25-DEC-1998; 98JP-0369093.
 XX 25-DEC-1997; 97JP-0369443.
 PR (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.
 DR WPI: 1999-554023/47.
 DR P-PSDB: AAY29336.
 PT New caspase-activating DNase - useful for eliminating cancer cells
 PS Claim 4; Page 14; 20pp; Japanese.
 CC The present sequence encodes mouse caspase activating DNase (CAD).
 CC CAD can be used for the removal of cancer cells.
 XX
 SO Sequence 1038 BP; 236 A; 293 C; 307 G; 202 T; 0 other;
 Alignment Scores:
 Pred. No.: 3,886-138 Length: 1038
 Score: 1416.00 Matches: 261
 Percent Similarity: 88.17% Conservative: 37
 Best local Similarity: 77.22% Mismatches: 38
 Query Match: 79.15% Indels: 2
 DB: 20 Gaps: 1
 US-09-748-451-2 (1-338) x AAZ21180 (1-1038)
 QY 1 MetLeuGlnLysProLysSerValIleuArgAlaLeuArgSerProArgLysPheGly 20
 |||||
 Db 10 GTGTCGCGCCACACCAATGCGTAACTTGAGAGCCCTACATAGCGCTGCACTTGGC 69
 |||||
 QY 21 ValIleArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 |||||
 Db 70 GTGCGCGCCGCGAGACTGCCAGAGCTGCTGCGTAAAGGGCTGCGCTCCACACTCCG 129
 |||||
 QY 41 GluArgGlySerArgLeuCysLeuTyrGlnAspLysThrGluLeuThrArgLysPhe 60
 |||||
 Db 130 ATGCCCGGTCCCGCGCTGTGCTGTACAGATGACGAGGAGTGCAGACGACTGCTTC 189
 |||||
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTrpGlnGly 80
 |||||
 Db 190 CCGGCGCTTCCCAACGAGCGTGAAGCTCTTATGCTACCGCTGCGAGACCTGCGATGC 249
 |||||
 QY 81 TyrValSerAspIleArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
 |||||
 Db 250 TATGTGAGTGAACATCACAGCTTCTCCAGTGTATGATGAGCCACATGCCGGGTCATC 309
 |||||
 QY 101 GlnAlaIleGlnGlnLeuLeuCysAspGlnGlnAlaProGlnAlaArgLeuAla 120
 |||||
 Db 310 CAGGCTGCAGGACGCTGTCTGATGAGCAGGCCCTCCTGAGAAAGCTCTGACC 369
 |||||
 QY 121 AspleuLeuHisAsnValSerGlnAsnIleAlaIleGluThrArgAlaGluAspPro 140
 |||||
 Db 370 GATCTTGTGATCATCGAGCCAGAACATATTAATCTCAAGACCCGGAGCAGACCCATCC 429
 |||||
 QY 141 TrpPheGluGlyLeuGlnSerArgPheGlnSerIleSerGlyTyrLeuArgTyrSerCys 160
 |||||
 Db 430 TGGTTGAAGGTTTGGAGTCGAGATCAGGAATTAAGCGGCTATGAGATACAGCTGT 489
 |||||
 QY 161 GluSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
 |||||
 Db 490 GAGAGTCGATCGAGGCTTACCTTAAGAGAGGTGAGCCCTTACCTTATGCTGATGAA 549
 |||||
 QY 181 GluAlaGlnGluGlnPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
 |||||
 Db 550 GCAGCTCAAGAGAGTACCTGCGAGTCTTGGCTCCATGTGCCAGAACTCAATCGGTG 609
 |||||

```

OY 201 GATYAAAGLYSEYTYRPHASPAAGLYALYSGLYSERATYLEUCYTHPRO 220
DB 610 CAGTACAAAGGACGATATTCGACAGAGGCGAGAGCCGCGCTCTGACATCCA 669
OY 221 GUGLYTPRPHESERCYSGINGLYPROPHASPMETASPSERCYLAUSERARGHISER 240
DB 670 GAAGATGCTTCTCTGCGACAGGCGCTTGACCTGGAAGCTGCTTCCACACTCC 729
OY 241 ILEAANPROTYRSEASNAARGGLUSERARGILELEPHESERTHTTPASNALEASPHIS 260
DB 730 ATCAACCCCTATGCAACAGAGAGAGCGGATCTCTTCACTGACACTGGAACCTGATCAT 789
OY 261 ILEILEGLLYLYSARGTHRIILEPROTHLEUVALIGLUALILELYSGLYGLNASP 280
DB 790 AATAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
OY 281 GUYATGGLYUVALASPTPGIUTYRPHETRYGLYLEULEUPHETHTSERGLUASNALEU 300
DB 844 GGGAGGAGAGTGAAGTGGAGAGTCTTACAGCTCTCTTACAGCTCTCTTACAGCTCT 903
OY 301 LEUVALHISILEVALCYSHALYSLYSLYTHRIHISLYLEUASNCYASPPROSERARG 320
DB 904 CTGATGACATGCGCTGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 963
OY 321 ILETYRLYSPROGLINTHARGLEULYSLYSLYGLNPROVALARGLYSARGLN 338
DB 964 ACTATCGGCTCTGACAGAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
RESULT 7
AA189481/c
ID AA189481 standard; cDNA; 379 BP.
AC AA189481;
AD 06-NOV-2001 (first entry)
DE Human polynucleotide SEQ ID NO 9541.
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
OS Homo sapiens.
PN WO200164835-A2.
OY 07-SEP-2001.
DB 26-FEB-2001; 2001WO-US04927.
PR 28-FEB-2000; 2000US-0515126.
XX 18-MAY-2000; 2000US-0577409.
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
DR WPI; 2001-514838/56.
XX P-PSDB; AAO09550.
PT Isolated nucleic acids and polypeptides, useful for preventing
PT disorders and treating e.g. leukaemia, inflammation and immune
XX disorders.
PS Claim 1; SEQ ID NO 9541; 1399pp + Sequence Listing; English.
CC The invention relates to human polynucleotides (AA189941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,

```

```

CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
SQ Sequence 379 BP; 88 A; 90 C; 97 G; 104 T; 0 other;

Alignment Scores:
Pred. No.: 1,07e-29 Length: 379
Score: 372.00 Matches: 76
Percent Similarity: 75.24% Conservative: 30
Best Local Similarity: 72.38% Mismatches: 10
Query Match: 20.79% Indels: 16
DB: 22 Gaps: 2

US-09-748-451-2 (1-338) x AA189481 (1-379).
OY 234 SERCYLAUSERARGHISERTILEASNPOTYRSEASNAARGLYSERARGILEUPHE 253
DB 341 TCTGCTCTGAGC-----CTCATG 324
OY 254 SETHTTRPASNLEUASPHISILEILEGLLYSLYSLYTHRIILEPROTHLEUVAL 273
DB 323 AGTCGCTGGGCGC-----TACAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 270
OY 274 GLUALALELYSGLYGLNASPGLYARGGLYUVALASPTPGIUTYRPHETRYGLYLEU 293
DB 269 GAAGCAATTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 210
OY 294 PHEHTHRSERGLUASNPOTYRSEASNAARGLYSERARGILEUPHE 313
DB 209 TTACCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 150
OY 314 LEUASNCYASPPROSERARGILEYRYSPTGINTHARGLEULYSLYSLYGLNPRO 333
DB 149 CTCACCTGTGACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 90
OY 334 VALARGLYSARGLN 338
DB 89 GTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 75
RESULT 8
ABL15179
ID ABL15179 standard; cDNA; 1353 BP.
AC ABL15179;
AD 26-MAR-2002 (first entry)
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 40019.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
OS Drosophila melanogaster.
PN WO200171042-A2.
OY 27-SEP-2001.
DB 23-MAR-2001; 2001WO-US09231.
PR 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
PA (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PMD, Myers EW;
XX

```



```

OY 86 ArgArgPheLeuSerAlaPheHisGlu-----ProGlnValGlyLeuIle 100
DB 2002 TGCTCTCTATTGATTCGATTTTGGATTCGAGAGATGCGACCAATCCCTTTGCTA 1943
OY 101 GlnAlaIaIaGlnIleLeuLeuCysAsp-----GlnGlnAlaProGlnArg---GlnArg 117
DB 1942 AAGGTGGCT---GACATTTTCTTACGATTTTATTCGACACATCCGGAATAATTCGCGCGC 1886
OY 118 LeuLeuAlaAspLeuLeuHisAsnValSerGlnAsnIleAlaIaIaGlnIle 134
DB 1885 ATGATTAACGCGAGTATGTAACACCAAAAGACGAGCGCTCGATTAACAGCAAGACCAAC 1826
OY 135 -----ArgAlaGlnAspProTProTProTProTProTProTProTProTProT 1766
DB 1825 CTCAGCTTAAAGGCGGACGACGTCGAGTTCACGGCGCGGAGGAGCGCTTCACCTCT 1766
OY 152 LysSerGlyTyrLeuArgTyrSerCysGlnSerArgIleArgSerTyrLeuArgGluVal 171
DB 1765 AAGGAGAGACGTATGCGCACACGTCGCGCACACGTCGCGAGGCTCTAC----- 1715
OY 172 SerSerTyrProSerThrValGlyAlaGlnIaGlnIaGlnIaGlnIaGlnIaGlnIa 188
DB 1714 -----TATAAGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1679
OY 189 -----ValLeuGlySerMetCysGlnArgLeuArgSer 199
DB 1678 TACCGCCAAATGCAAGCCGACGAGTAAACCTGCTGCGAGAGAGAGAGAGAGAGAGAG 1619
OY 200 Met-----GlnTyrAsnGlySerTyrPheAspArgGlyAlaLys----- 212
DB 1618 CTGCTCATCGGTTGCGACTTTTCTCCATGATGTTTACCGGAAATTCGACCAAAAGCAT 1559
OY 212 ----- 212
DB 1558 GAATTCCTGAAGCAGACACTTGGCGGACGAGAAACAGACGCTGCGAGATACCCAGCAAG 1499
OY 213 -----GlyGly 214
DB 1498 AGACTGAGGAGAGTATGAGGAGTACACCAAGGAAATCTGATCTCGAGAAATGGTCC 1439
OY 215 SerArgLeuCysThrProGlnGlyTyrPheSerCysGlnGlyProPhe-----AspMet 232
DB 1438 ACTTCCTGCTGCTCGGACTTGGGACTTCTATTCGACAGGCTCTTATTCGAGAAATGCGC 1379
OY 233 AspSerCysLeuSerArgHisSerIleAsnProTyrSerAsnArgGluSerArgIleLeu 252
DB 1378 AACAGCTGCTGAAGCAGACACATTAATCCGTAACGCTGCGAGAAATCTCATTTTG 1319
OY 253 PheSerThrTrpAsnLeuAspHisIleIleGlnLysArgThrIleIleProThrLeu 272
DB 1318 TTTTCAGGTCTGGAATCTGGACCAACAAATCGAACTGTGCCGACAAATCTTCCGCACTT 1259
OY 273 ValGlnAlaIleLysGlu-----GlnAspGly 281
DB 1258 GTTGCAGAAATGGAAGAACTTGTGAGTACCCGACACCAAGTGTTCATTAATCAATAGAG 1199
OY 282 ArgGluValAsp-----TrpGluTyrPheTyrGlyLeuLeuPheThrSerGluAsn 298
DB 1198 CAAGTGTGATATCTCAGTACTGAGTACTGACTTCTAGAA---AATATCTCCCTTAAGAAAT 1142
OY 299 LeuIleValIleHisIleValCysHisLysLys-----ThrThrHis 312
DB 1141 CTCAAACTAGTGCACATTTGCTGTGACAGAAAGCGGAGCGGCAAAACCTTTCGAATGGT 1082
OY 313 LysLeuAsnCysAspProSerArgIleTyrLys 323
DB 1081 CGCCTACTGTCTCCGACTGTCTCATGAGTATCGC 1049

```

```

XX 14-AUG-2002 (first entry)
DE Human CDNA differentially expressed in granulocytic cells #1334.
DE
XX Human; ss; granulocytic cell; DNA chip; bacterial infection;
XX viral infection; parasitic infection; protozoal infection;
XX fungal infection; sterile inflammatory disease; psoriasis;
XX rheumatoid arthritis; glomerulonephritis; asthma; thrombosis;
XX cardiac reperfusion injury; renal reperfusion injury; ARDS;
XX adult respiratory distress syndrome; inflammatory bowel disease;
XX granulocyte activation; chronic inflammation; allergy.
XX Homo sapiens.
XX WO200228999-A2.
XX 11-APR-2002.
XX 03-OCt-2001; 2001WO-US30821.
XX 03-OCt-2000; 2000US-237189P.
XX (GENE-) GENE LOGIC INC.
XX Beazer-Barclay Y, Weissman SM, Yamaga S, Vockley J;
XX WPI; 2002-435328/46.
XX
XX Detecting granulocyte activation by detecting differential expression
XX of genes associated with granulocyte activation, which serves as
XX diagnostic markers that is useful for monitoring disease states and
XX drug toxicity
XX
XX Claim 1; SEQ ID NO 1334; 114pp; English.
XX
XX The invention relates to detecting (M1) granulocyte (GC) activation
XX (GCA), by detecting the level of expression of gene(s) (Gs) identified by
XX DNA chip analysis as given in the specification, and comparing
XX the expression level to an expression level in an unactivated
XX GC, where differential expression of Gs is indicative of GCA.
XX Also included are modulating (M2) Gs by contacting GC with an agent
XX that alters the expression of at least one gene in Gs; (2) screening (M3)
XX for an agent capable of modulating GCA or an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease using the
XX gene expression profile; (3) detecting (M4) an inflammation (especially
XX chronic) in a tissue, an allergic response in a subject, exposure of a
XX subject to a pathogen or sterile inflammatory disease, by detecting the
XX level of expression in a sample of the tissue of gene(s) from Gs, where
XX the level of expression of the gene is indicative of inflammation;
XX (4) treating (M5) an inflammation (especially chronic) or in a tissue,
XX an allergic response in a subject, exposure of a subject to a pathogen
XX or sterile inflammatory disease, by contacting a tissue having
XX inflammation with an agent that modulates the expression of gene(s)
XX from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for
XX modulating Gs; M3 is useful for screening an agent capable of modulating
XX GCA preferably in an inflammation in a tissue; M4 is useful for
XX detecting an inflammation (especially chronic) in a tissue, an allergic
XX response in a subject, exposure of a subject to a pathogen or sterile
XX inflammatory disease (e.g. psoriasis, rheumatoid arthritis,
XX glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal
XX reperfusion injury, ARDS, adult respiratory distress syndrome,
XX inflammatory bowel disease, Crohn's disease, ulcerative colitis,
XX periodontal disease, also bacterial infection, viral infection,
XX parasitic infection, protozoal infection, fungal infection and M5 is
XX useful for treating one of the above conditions. The present
XX sequence represents a gene differentially expressed in granulocytes.
XX Note: The sequence data for this patent did not form part
XX of the printed specification, but was obtained in electronic
XX format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

```


CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.

XX SQ Sequence 782 BP; 200 A; 205 C; 198 G; 179 T; 0 other;

Alignment Scores:

Pred. No.:	0.238	Length:	782
Score:	105.00	Matches:	46
Percent Similarity:	37.88%	Conservative:	29
Best Local Similarity:	23.23%	Mismatches:	69
Query Match:	5.87%	Indels:	54
DB:	21	Gaps:	11

US-09-748-451-2 (1-338) x AAA01540 (1-782)

```
168 LeuArgGluValSerSerTyrProSerThrValGluAlaGluGluPhe-- 186
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
658 CTTCAACAACTGATGATTTGCCCTTGAGATTCGCCCAAGCCGACAGCAAAATTTAG 599
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
187 -----LeuArgValLeuGly----- 191
598 AAACAGGTACGCCCTTCTGCACAGAGTTGTTGCCACTAGTCCGCGGTTCTGACA 539
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
192 -----SerMetCysGlnArg-----LeuArgSerMetGlnTyr 202
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
538 GGGCGCTGATGTTATTTTCAGAACCCATTAAGTGGTATGGCTTCGACGGGACAGATT 479
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
203 AsnGlySerTyrPheAspArgGlyAlaLysGlyLysSerArgLeuCysThrProGluGly 222
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
478 ---GGACATTGGTTCGACCTCAGGCGGT----- 455
223 TrpPheSerCysGlnGlyProPheAspMetAspSerCysLeuSer---ArgHisSerIle 241
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
454 -----GGCCCAATA---CTTCAAGTTGTCTGTAGCTTGGCGCGCAATG 413
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
242 AsnProTyrSerAsnArgGluSerArgIleLeuPheSerThrTyrAsnLeuAspHisIle 261
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
412 AATCCCTTATCGAAGAACCGAACAGGCTTTGGCAGACAGGAATGTTTCATCAACATATT 353
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
262 IleGlu-----LysLysArgThrIleIleProThrLeuValGluAla 275
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
352 GCGAGAGCTCTCCACACAGACAGACAGCGCGGTTGCGGTTCTGGCTCGGATT 293
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
276 IleLysGluGlnAspArgGluValAspTrpGluTyrPheTyrGlyLeuLeuPheThr 295
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
292 CGGGGGGACGACACCGCATCGCGCATCTACGCCAC-----CTTGGCTTGCTTTATCA 239
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
296 SerGluAsnLeuLeuVal---HisIleValCysHisLys---LysThrThrHisLys 313
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
238 AGTAGCGGTTTTCACAGCATTCGCCATAGTTGTGCTAGCATATTCGTGCAGCTGGCCGC 179
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
314 LeuAsnGlyAspProSerArgIleTyrLysProGlnThrArgLeuLysArgLys 331
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
178 TTACGGGTAAACCTGTAACGACGCCCTGCGCCGACACAGAGATGCCGAAGCGA 125
    |||  :::::  |||  :::  |||  :::::  |||  :::::  |||
```

RESULT 15

AAA09307
ID AAA09307 standard; DNA; 1766 BP.

AAA09307;

10-AUG-2000 (first entry)

Human vimentin coding sequence.

XX NIK1 Interacting Protein; vimentin; protein complex; cytoskeletal;
KW antiviral; neuroprotective; cardiant; ss.
XX

OS Homo sapiens.

XX Key Location/Qualifiers

XX Key 44..1444

XX CDS /tag= a

XX W0200020448..42.

XX 13-APR-2000.

XX 06-OCT-1999; 99WO-US23314.

XX 06-OCT-1998; 98US-0167206.

XX (CURA-) CURAGEN CORP.

XX Nandabalan K, Schulz VP, Yang M;

XX WPI, 2000-303742/26.

XX P-PSDB; AAY92335.

XX Example 1; Page 142-145; 172pp; English.

AAA09303-09 were isolated in a modified yeast two hybrid system using
NIK1 protein as "bait". These are known sequences which encode NIK1
interacting proteins. The invention concerns purified complexes of a
NIK1 protein and a NIK1 protein-interacting protein, where the
interacting protein is chosen from TrkA, protein phosphatase 1alpha,
14-3-3epsilon, alpha-tropomyosin, vimentin, p0071, Inl-1, IP-1, IP-2,
IP-3, IP-4, or IP-5. NIK1 (also referred to as Nek2) is a human
homologue of the Aspergillus nidulans mitotic regulator, NIMA kinase.
NIK1 is a serine/threonine-specific kinase and is thought to play a key
role in cell-cycle events leading to the onset of mitosis. The
complexes, their derivatives and NIK1 or NIK1-IP protein and DNA
sequences, etc. are useful for treating or preventing a disease or
disorder involving aberrant levels of the complex or protein. Such
disorders include cancer, hyperproliferative disorders, such
as neurodegenerative disorders, cardiomyopathies, viral infections and
metabolic disorders.

XX SQ Sequence 1766 BP; 479 A; 482 C; 449 G; 356 T; 0 other;

Alignment Scores:

Pred. No.:	1.4	Length:	1766
Score:	102.50	Matches:	77
Percent Similarity:	34.42%	Conservative:	50
Best Local Similarity:	20.87%	Mismatches:	128
Query Match:	5.73%	Indels:	115
DB:	21	Gaps:	16

US-09-748-451-2 (1-338) x AAA09307 (1-1766)

```
4 LysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyValAlaGly 23
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
546 AAGCCCGCGTCC---AGGTGGAGCGCGCAACACCTGGCGGAGACATCATCGCGCGGG 602
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
24 ArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuProGluArgGly 43
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
603 AGAATTGACAGG----- 614
44 SerArgLeuCysLeuTyrGluAspGlyThrGluLeuThrGluAspTyrPheProSerVal 63
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
615 ---AGGAGATGCTTCAACAGACAG-GAAGCCGAAACACCTGCATATTTTCACAGAGAT 670
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
64 ProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaTrpGlnGlyTyrValSer 83
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
671 GTTACAAATGCGTCTCTGGCAGCGTCTGACCTTGACCAAGCAAGTCAATCTTCCACAGA 730
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
```


CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 7545 BP; 2136 A; 1578 C; 1563 G; 2268 T; 0 other.

Alignment Scores:	
Pred. No.:	33.6
Score:	98.10
Percent Similarity:	34.04%
Best Local Similarity:	19.15%
Query Match:	5.48%
DB:	23
Length:	7545
Matches:	72
Conservative:	56
Mismatches:	132.12
Indels:	117
Gaps:	16

US-09-748-451-2 (1-338) x ABL14002 (1-7545)

[illegible][illegible]

```
Search completed: May 26, 2003, 16:31:19
Job time : 729 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 15:24:28 ; Search time 63 Seconds
(without alignments)
1645.345 Million cell updates/sec

Title: US-09-748-451-2
Perfect score: 1789
Sequence: 1 MLOPKSVKRLAIRSPKRC.....SKIKPQRLKRPVKKRQ 338

ling table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 44362 seqs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-xlh
-O-/cgn2_1/uspto.spool/us09748451/unnat.21052003.153830.16303/app.query.fasta_1.519
-DB-issued Patents: NA -OFAST-fastap -SUFIX-trn1 -MINMATCH=0.1 -LOOPL=0
-LIST-45 -DUALS-bits -SPAR=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdl
-MODE-LOCAL -OFAST-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09748451.ecgn1.1.40.ernnat.21052003.153830.16303 -NCPU=6 -ICPU=3
-NO_MMAP -IAREOUTERY -NEG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -GAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -GAPOP=10 -GAPEXT=0.5 -DLOP=6 -DEEXT=7

Applied with only

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1789	100.0	2839	4	US-09-061-702-1
2	101.5	5.7	1209	1	US-08-553-703A-5
3	101.5	5.7	1209	1	US-08-553-703A-9
4	101.5	5.7	1209	2	US-09-006-021-5
5	101.5	5.7	1209	2	US-09-006-021-9
6	96	5.4	2729	1	US-07-938-782A-1
7	96	5.4	2729	1	US-08-630-524-1
8	96	5.4	2729	3	PCT-US93-08131-1
9	94	5.3	2143	4	US-09-083-023-21
10	91	5.1	1167	4	US-09-083-023-19
11	89.5	5.0	2301	4	US-09-083-198B-8
12	89	5.0	1689	4	US-09-061-702-3

13	88.5	4.9	9633	6	US-09-069-023-13	Patent No. 5223423
14	88	4.9	1049	4	US-08-794-158-1	Sequence 13, Appl
15	88	4.9	2637	4	US-08-794-158-1	Sequence 1, Appl
16	88	4.9	2637	4	US-08-794-158-3	Sequence 3, Appl
17	86.5	4.8	4322	2	US-08-537-342-1	Sequence 1, Appl
18	86	4.8	1114	4	US-09-069-023-17	Sequence 17, Appl
19	85.5	4.8	4403765	4	US-09-103-840A-2	Sequence 2, Appl
20	85.5	4.8	4411529	1	US-08-707-792A-2	Sequence 1, Appl
21	85	4.8	1155	1	US-08-707-792A-2	Sequence 2, Appl
22	85	4.8	1155	1	US-09-329-884-13	Sequence 13, Appl
23	84.5	4.7	1124	3	US-08-945-056-1	Sequence 4, Appl
24	84.5	4.7	1124	3	US-08-945-056-4	Sequence 14, Appl
25	84.5	4.7	4201	3	US-08-799-464A-14	Sequence 14, Appl
26	84.5	4.7	15101	5	PCT-US95-09927-14	Sequence 1, Appl
27	84.5	4.7	15101	5	US-08-157-005-1	Sequence 1, Appl
28	84.5	4.7	15108	4	US-08-747-863-1	Sequence 1, Appl
29	84.5	4.7	15108	4	US-09-565-864-1	Sequence 1, Appl
30	84.5	4.7	15108	4	US-08-244-189-1	Sequence 1, Appl
31	84	4.7	4287	1	US-08-306-691B-53	Sequence 53, Appl
32	84	4.7	20235	1	US-07-642-734C-3	Sequence 3, Appl
33	84	4.7	20235	3	US-08-439-009A-3	Sequence 3, Appl
34	83.5	4.7	2680	2	US-08-533-306A-3	Sequence 5, Appl
35	83.5	4.7	2680	2	US-08-742-923A-5	Sequence 5, Appl
36	83.5	4.7	4403765	4	US-09-103-840A-2	Sequence 2, Appl
37	83.5	4.7	4411529	4	US-08-873-093-2	Sequence 2, Appl
38	83.5	4.6	2268	2	US-09-206-646-2	Sequence 2, Appl
39	82.5	4.6	2268	2	US-08-659-251-14	Sequence 14, Appl
40	82.5	4.6	3168	4	US-09-256-490-14	Sequence 14, Appl
41	82.5	4.6	3168	5	PCT-US96-11445-14	Sequence 14, Appl
42	82.5	4.6	3168	5	US-08-659-251-1	Sequence 1, Appl
43	82.5	4.6	10163	4	US-09-256-490-1	Sequence 1, Appl
44	82.5	4.6	10163	4		
45	82.5	4.6	10163	4		

ALIGNMENTS

US-09-061-702-1
Sequence 1, Application US/09061702
Patent No. 6165737
GENERAL INFORMATION:
APPLICANT: Wang, Xiaodong
APPLICANT: Liu, Xuesong
TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN
TITLE OF INVENTION: APOPTOSIS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/061,702
FILING DATE: Concurrently Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: McMillian, Nabesla R.
REGISTRATION NUMBER: P-43,363
REFERENCE/DOCKET NUMBER: UTSD:546
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512)418-3000
TELEFAX: (512)474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2839 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-061-702-1

Alignment Scores:

Pred. No.:	6,556-208	Length:	2839
Score:	1789.00	Matches:	338
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-748-451-2 (1-338) x US-09-061-702-1 (1-2839)

```

QY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
DB 132 ATGTCTCAGAGAGCCAGAGCGTGAAGCTGCGGCCCTGGGACCGGAGAGATTGGC 191
QY 21 ValAlaGlyArgSerGlyGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
DB 192 GTGGCTGGCCGGAGCTCCAGAGCTGCTGCGCAAGGGCTGTCTCCCTCCAGCTCCCT 251
QY 41 GluArgGlySerArgLeuGlyCysLeuTyrGluAspGlyThrGluLeuThrGluAspTyrPhe 60
DB 252 GAGCGCGGTTCCCGGCTGTGCTGTACGAGAGATGCGAGAGCTGACGAGGATTACTTC 311
QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaArgPheGlnGly 80
DB 312 CCCAGTGTCCCGCAACAGCGCGAGCTGTGCTGCTGCTGCGGCCGAGGCTGGCAGAGGC 371
QY 81 TyrValSerAspLeuArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
DB 372 TATGTGAGCCAGATCAGCGCTTCTCCAGATTCACAGCCACAGGGGGCTTCATC 431
QY 101 GlnAlaAlaGlnGlnLeuLeuGlyCysAspGlnGlnAlaProGlnArgGlnArgLeuAla 120
DB 432 CAGGCGCCCGCAACAGCTGCTGTGTGTGATGACAGGCCCCACAGAGCGAGGCTGTGCT 491
QY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspProPro 140
DB 492 GACCTCTCTGCACAACTCAGCAGACCATGCGGCGCCAGAGCGGGCTGAGAGCCGCCG 551
QY 141 TrpPheGlnGlnLeuGlnSerArgPheGlnSerLysSerLysTyrLeuArgTyrSerCys 160
DB 552 TGGTTTGAAGGCTTGAGTCCCATTTCAAGACAAAGCTGCGATTCGACATACAGCTGT 611
QY 161 GluSerArgLysLeuArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
DB 612 GAGAGCGCGATCCGGAATTACCTGAGGAGGTGAGCTTCACTCCACAGTGGTGGC 671
QY 181 GluAlaGlnGlnGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
DB 672 GAGGCTCAGAGGAAATCTGCGGGTCTCGGCTCATGTCAGAGAGCTCCGGTCCATG 731
QY 201 GlnTyrAsnLysSerTyrPheAspArgGlyAlaLysGlyLysSerArgLeuGlyCysThrPro 220
DB 732 CAGTACATAGGCGAGCTTCTGACAGAGGAGCCAAAGGGCGGACCCCTCTGCACACCG 791
QY 221 GlnGlyTyrPheSerGlyGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
DB 792 GAAAGCGGTTCTCTGCGAGGCTCTTTCATGACATGACAGCTTCTTATCAAGACATCC 851
QY 241 IleAsnProTyrSerAsnArgLysSerArgIleLeuPheSerThrTrpAsnLeuAspHis 260
DB 852 ATCAACCCCTTACAGTACAGAGAGAGAGATCCTCTTACAGACCTGGAACCTGGATCAC 911
QY 261 IleIleGlnLysArgThrIleIleProThrLeuValAlaIleLysGlnGluAsp 280
DB 912 ATATATGAAAAAGAACCCATCATCTTACACTGTGTGAAGCAATTAAGAAACAGAT 971
QY 281 GlnArgGluValAspTyrGlnTyrPheTyrGlyLeuLeuPheThrSerGlnAsnLeuLys 300

```

```

DB 972 GGAAGAGAGATGCACTGGGAGTATTTTATGAGCCCTGTTTAACTTCAGAACCTTAAAA 1031
QY 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSerArg 320
DB 1032 CTAGTGCACATTTGCTCTGCTCCATTAAGAAAAACCAACGCTCAACTGTGACCCGAGAGA 1091
QY 321 IleTyrLysProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
DB 1092 ATCTAACAAACCCAGACAGGTTGAAGCGGAGAGAGCTGTGGGAAACGCCAG 1145

RESULT 2
US-08-553-703A-5
Sequence 5, Application US/08553703A
Patent No. 5795767
GENERAL INFORMATION:
APPLICANT: MARU, ISAFUMI
APPLICANT: OHTA, YASUHIRO
APPLICANT: TSUKADA, YOJI
TITLE OF INVENTION: PRIMERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/553,703A
FILING DATE: 30-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SRO ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1209 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-553-703A-5

Alignment Scores:
Pred. No.: 0.0142 Length: 1209
Score: 101.50 Matches: 79
Percent Similarity: 38.97% Conservative: 34
Best Local Similarity: 27.24% Mismatches: 98
Query Match: 5.67% Indels: 79
DB: 1 Gaps: 16

US-09-748-451-2 (1-338) x US-08-553-703A-5 (1-1209)
QY 31 ArgLysGlyCysLeuArgPheGlnLeuProGlnArgGlySerArg-----LeuGly 47
DB 63 CGTATGCGCTTTCTGCGTGGAGCACTCCAGATCGGAGACAGGGGCTTCTTCAACGTG 122
QY 47 sleuTyrGluAspGlyThrGluLeuThrGluAspTyrPheProSerValProAspAsnAl 67
DB 123 CCGTGGCGCGGACGCGG-----CGGTGTATGACGACCT 155

```

Tue May 27 09:31:56 2003

us-09-748-451-2.rni

Page 3

OY	67	aclueleuallenucllylneuclylaalatrpglnclnytyrvalsetarpiileatg	87
Dd	156	cAAGTACTGCTGGCTGCAGGAGGACAGGTGTGATG-----TACTGCCCTGTACC	209
OY	87	gheleusearlaphetisgluproglnvalglyleuileglnalaalaglnl----	105
Dd	210	CAGGTTCAGCGCCTCCACCGCCTGAG-----CTTGTGATGCGCGCTAAAGCAGGCGG	263
OY	106	-----leuLucyAspluglnalalaproglinalvarglnAlvgleuleualalaplue	123
Dd	264	CGAATTTTGCTGGCCCATGCCCAGGTGGCACCTCTGAAAACAATGTGCTTTGTGCT	323
OY	123	uhls-----AsvalSerglnashllealalaglU-----Th	134
Dd	324	GACGGGGGACGGCCGGCCGTCAGAAGTCCAGCGAGCATTTCAAGTAAGTCTTCTAC	383
OY	134	rArglaIguaspproPrtPpheglUglyleuGluseArpPhglnSerLyssergl	154
Dd	384	CATGGCCATGAACAGCTGTGGAGGATACGGCGAGCAGCGTACAGAGGAAACCG--	441
OY	154	yrryleuarqryserysercysgluseArpIlavrgserTyrlleuarqglvalsetse-	173
Dd	442	-----ctGACACATGATGATGATGCATGTGTCACATGGGTCCAGAGAACCTCTGG	491
OY	174	-----TyProserThrValglyAlaglnalnclnluU-----	185
Dd	492	GCTGGGCGGCGCCACAGCTCCCGGGGGCTGCGCATGCGCATGCGCATGCCATGAT	551
OY	186	-pheleuarqyalneu-glysermetCysglnarqlleuarqsermetglntryanelys	205
Dd	552	GCTGTGTGCGCTGGTGTGACACTCGGGGAAGAGACAGAGACCTGGCAGCGCGTAC	611
OY	205	eTrYPheaspArqglAlalysglilyserAlgleuCyThr-ProdlnclYTrpPhe	224
Dd	612	GCAGCTGGGGGACAGTGCGCTGCGTAGAGATCC---TCACACACAGTCCAGAGGATGCA	668
OY	225	SerCYsglnIgyPro-Phe-----AspmetAspSerCysLeuSerArgH1	239
Dd	669	GGCTGTGCTGGAGATGTCTCGAAGATGCGAGAGAACTTTGCTGCTCGTGGAGACA	728
OY	239	sSerlleasnPro-----TyrsertAsnAr	247
Dd	729	CCAG-----AACCCAGGCGACGCGGTGAAGTGGCTTCCTGCTCCGCCACAGAGCGC	785
OY	247	g-----gluseArgllealePhe-----	253
Dd	786	GAGCGGTGACCCCAAACCTCGACGCCACGATCATGTGAACGCTTCCTACTGCCCTTCG	845
Dd	846	CTCCGATGGGACGCTGATCAC	867
 RESULT 3 DS-08-553-703A-9 Sequence 9, Application US/08553703A Patent No. 5795767 GENERAL INFORMATION: APPLICANT: MARU, ISAFUMI APPLICANT: OHATA, YASUHIRO APPLICANT: TSUKADA, YOJI TITLE OF INVENTION: EPIMERASE NUMBER OF SEQUENCES: 9 CORRESPONDENCE ADDRESS: ADDRESSEE: Knobbie, Martens, Olson & Bear STREET: 620 Newport Center Drive 16th Floor CITY: Newport Beach STATE: CA U.S.A. ZIP: 92660 COMPUTER READABLE FORM: MEDIUM TYPE: Diskette COMPUTER: IBM Compatible			

```

? OPERATING SYSTEM : DOS
? SOFTWARE : FastSeq Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER : US/08/553_703A
? FILING DATE : 30-NOV-1995
? CLASSIFICATION : 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME : Altman, Daniel E
? REGISTRATION NUMBER : 34,115
? REFERENCE/DOCKET NUMBER:
? TELECOMMUNICATION INFORMATION:
? TELEPHONE : 714-760-0404
? TELEFAX : 714-760-9502
? TELEX:
? INFORMATION FOR SEQ ID NO: 9:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1209 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? FEATURE:
? NAME/KEY: Coding Sequence
? LOCATION: 1..1206
? OTHER INFORMATION:
US-08-553-703A-9

Alignment Scores:
Pred. No.: 0.0142 Length: 1209
Score: 101.50 Matches: 79
Percent Similarity: 38.97% Conservative: 34
Best Local Similarity: 27.24% Mismatches: 98
Query Match: 5.67% Indels: 79
Gaps: 16

US-09-748-451-2 (1-338) x US-08-553-703A-9 (1-1209)
QY 31 ArgLysGIcYSLeuArGPheGLneUProGLuaRgLYSerArg-----LeucY 47
Db 63 CGTGTGGCTTTTCGTGCCTGAGCAGCTCCCGACATCGGAGCACGGGGCTTCTTCAGTG 122
QY 47 sLeuTYrGLuaSPeLIyThrGLueUThluASPTyr-PheProSeVaIProASPanaI 67
Db 123 CCTGGGCCGCCACGG-----CGGGTGTATGACGACT 155
QY 67 aGLueValLeuLeuThrLeuGLyGLAlaTRPGLngLYTyR-VaISerAsPIleARgAr 87
Db 156 CAAGTACCTCTGCCTGAGGAGGAGGAGAGGTGTGGATG-----TACTGTCCGCTGTACC 209
QY 87 gPheLeuSerLaIaPheHISGLUPROGLIVaIGLYLeuILeGLIALaIaGLInGLn---- 105
Db 210 CAA GTTGAGCGCTTCCACGCCCTTAG-----CTTCTGTATGCGGCTAAAGCAGGGGG 263
QY 106 -----LeueneuYaSPoJLuMaLProlnaRGlnARyLeuLeuLaIAAPLeuLe 123
Db 264 GCAATTTTTGTGGCCCATCGCCGAGTGEGCACCTCCGAANAANAOTGTGCTTGCTGT 323
QY 123 uHis-----AsnyVaSerLaSPnIIeAlaIaGLu-----Th 134
Db 324 GAACGGGGAGCGCCGCGCGCTCAAGTGTGACGAGAATCTTTAGTGAATGCTTTCAC 383
QY 134 rArGaIaGLuaSProProTrIPheGLIngLYLeuGLuSerArYPheGLInSerLYSergI 154
Db 384 CATGTCCATGAACAAGCTGTGGAGGTGTACGCGCGAGCAGGTATCCAGCGAAGCG-- 441
QY 154 yTYrLeuARgTYrSerCySGlUsErArGLIEarSYerTYrLeuARGLUaISerSer-- 173
Db 442 -----GTGACATGATGATGATNGATGTGCATGGGTGCGAGAGAACCCCTCTGG 491
QY 174 -----TYrProSerThrVaIGLYaGLaIaGLInGLn----- 185

```


Tue May 27 09:31:56 2003

us-09-748-451-2.rn1

Page 5

Patent No. 5994105
GENERAL INFORMATION:
APPLICANT: MARU, ISAFUHI
APPLICANT: OHYA, YASUHIRO
APPLICANT: TSUKADA, YOJI
TITLE OF INVENTION: EPIMERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA U.S.A.
ZIP: 92660

COUNTRY: U.S.A.
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/006,021
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/553,703
FILING DATE: 30-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1209 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1206
OTHER INFORMATION:
US-09-006-021-9

Alignment Scores:
Seq. No.: 0.0142 Length: 1209
Percent Similarity: 101.50 Matches: 79
Best Local Similarity: 38.97% Conservative: 94
Query Match: 27.24% Mismatches: 38
DB: 5.67% Indels: 79
Gaps: 16

US-09-748-451-2 (1-338) x US-09-006-021-9 (1-1209)
QY 31 ATGTGAGTCTGATGCGTTCGGTGACGACCTCCGCAGATCGGAGGGGCTTTCCACACTG Leucy 47
Db CGTAGATGCTTTCGTCGCGGAGCACCTCCGCAGATCGGAGGGGCTTTCCACACTG 122
QY 47 sileuTYrGLuaSPgiYmThrGLueurThrGLuaSPryPrheProSerValProAspaNAL 67
Db 123 CCTGGGCCGCGACGGG-----CGGCTGTATGACGACT 155
QY 67 acGUeuValleuLeuThrLeuGLyGlnAlATppGInGLyTYrValSerApIleaAr 87
Db 156 CAAATTTTGTGCGCATGCCCGAGTGCAGACCTCTCGAAGGAAGATGCGCTTTGCT 323
QY 87 pHeuLeuSerAlaPheHISglUpProGInValGLyLeuILeglaIAalagInLn----- 105
Db 210 CAAGCTTGAGCGCTTCACGCGCCCTGAG-----CTTCTGATGCGGCTTAAGCAGGGGG 263

QY 106 -----LeuLeuCysASPgInGLuaLProGInARgSLInARgLeuALeuaSPleu 125
Db 264 CGAATTTTGTGCGCATGCCCGAGTGCAGACCTCTCGAAGGAAGATGCGCTTTGCT 323
QY 123 uHLs-----AsnValSerGlnAnILeAlaLaGLu-----Th 134
Db 324 GAGCGCGGAGCGCGCGCCGCCCAAGGTGCAGAGGAGATCTTCAGTGAAGTCTTCAC 388
QY 134 rAlaGLaGLuaSProPOTrrPPheGLyGLyLeuGLuSerArIpheGlnSerTySergl 154
Db 384 CATGGCATGAACAGACTGTGAGGTGACGCGGAGCGACGCTACCAAGGANAAGC-- 441
QY 154 yTYrLeuArGYrSerCysGLuSerArIllaRgseryTYrLeuArgJuaLSer-- 173
Db 442 -----GTGACATGATGATGCAGATCTGCACTGGATGGTCCGAGAGACCCTCTGG 491
QY 174 -----TYrProSerThrValaGLyAlaGLuaLaGLInGLu----- 185
Db 492 GCTGGCGCGCGCCCGACCTCCCGGGCGCGCTGCGCTCGAGTCCATGACGATCCCAGAT 551
QY 186 -pheuArGuValleu-GlyserMeTySGlnARgLeuArGserMeGlNyTyranGLys 205
Db 552 GCTGCTGTGCTGTGGTGGAGACACTCGGAGAGAGAGACAGACTGGAGCGCGCTACGC 611
QY 205 eTyrPhesAprTrgLYalaLySGlySerArGLueCySrThr-ProGLuGLYTyrPhe 224
Db 612 GCAGCTGGGACACTGTGTGGTGTGGAAGATGGCAGAGAACCTTTCGCTGCTGGAGACA 666
QY 225 SerCYsGLNgLPPro-Phe-----AspMetLaPsetCYsLeuSerArGHl 239
Db 669 GGCTGTGCTGAGAAATGTGTGGAAAGATGGCAGAGAACCTTTCGCTGCTGGAGACA 722
QY 239 aserILEanSPo-----TYrSerAnAr 247
Db 729 CCAG---AACCAAGCGACACGGCGTGGAAAGCTGGTTCCTGCTCCGACACGACGG 785
QY 247 g-----gluSerArGlleuPhe----- 253
Db 786 GAGCGGTGACGCCAANCTTCAGACCCAGCTGATGACAGCTTCTACTGCTTCCG 845
QY 254 -SerThrTPasInLeuASPHls 260
Db 846 CTCGGATGGAGACGCTGATCAC 867

RESULT 6
US-07-938-782A-1
Sequence 1, Application US/07938782A
Patent No. 5525513
GENERAL INFORMATION:
APPLICANT: Chen, Jane J.
TITLE OF INVENTION: DNA Encoding the Heme-Regulated
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Papst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/938,782A
FILING DATE: 31-AUG-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Papst, Patrea L.

REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: W0101CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404)873-8794
 TELEFAX: (404)873-8795
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2729 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 ORGANISM: Rabbit
 CELL TYPE: Reticulocytes
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 113..2149
 OTHER INFORMATION: /note= "Expression of HRI
 OTHER INFORMATION: mRNA in Human erythroid cells, using as the
 OTHER INFORMATION: probe rabbit HRI cDNA from nucleotides 113
 OTHER INFORMATION: to 2149."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 229..249
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of human HRI cDNA sequence"
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 543..560
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of human HRI cDNA sequence"
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 448..468
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of human HRI cDNA sequence"
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1009..1031
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of a human HRI cDNA sequence"
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 PUBLICATION INFORMATION:
 AUTHORS: Chen, Jane J.
 AUTHORS: London, Irving M.
 TITLE: Cloning of the cDNA of the heme-
 TITLE: regulated eukaryotic initiation factor
 TITLE: 2alpha (elf-2alpha) kinase of rabbit reticulocytes:
 TITLE: Homology to yeast GCN2 protein kinase and human
 TITLE: double-stranded-RNA-dependent
 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
 VOLUME: 88
 PAGES: 7729-7733
 DATE: September-1991
 US-07-938-782A-1

Alignment Scores:

Pred. No.: 0.247 Length: 2729
 Score: 96.00 Matches: 62
 Percent Similarity: 35.43% Conservative: 45
 Best Local Similarity: 20.53% Mismatches: 121
 Query Match: 5.37% Indels: 74
 Gaps: 13

US-09-748-451-2 (1-338) x US-07-938-782A-1 (1-2729)

QY 4 LysProLysSerValLysLeuAlaLeuArgSerProArgLysPheGlyValAlaGly 23
 DB 1427 AAGCTGAGTAAATATTTCTCATGCTCTGATGACAAAGGAAATAGAGACTTGTG 1486
 QY 24 ArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPhe 38
 DB 1487 CTGGCTGGCGGCACATCATCCAGAAATGCGGCCGCCACACAGAAAGGGGAGAGA 1546
 QY 39 LeuProGluArgGlySerArgLeu-----CysLeuTyrgLysArgGlyThrGluLeuThr 56
 DB 1547 GCACCCACACACACTTCCCGAGTGGGACCGTGTGTACCGCTCGCCCGACCACTTGGA 1606
 QY 57 GluAspTyrgPheProSerValProAsp-----AsnAlaGluLeuValLeuLeuThrLeu 74
 DB 1607 GCATCGGAGTATGATGCCAAGTACAGATGATGACGCTCGCGCGATCTGTGACCTC 1666
 QY 75 GlyAlaIleArgPheGlnGlyValLeuSerAspLeuArgLysPheLeuSerAlaPheHisGlu 94
 DB 1667 -----TTCCAGCCCTTCGGGACAGAGATGGAGCGG--GCAGAGTCTTGACGGGC 1714
 QY 95 ProGlnValGlyLeuIleGlnAlaIleGlnIleLeuLeuCysAspGluAlaProGln 114
 DB 1715 GTGGAGCTGGCGGCATACCCGAGTCCCTGAGTAAAGGTGCCCGGGCGGACCAATAC 1774
 QY 115 ArgGlnArgLeuLeuAlaAspLeuLeuHisAsnValSerGlnAsnIleAlaIleGluThr 134
 DB 1775 GTCCAGCTGCTGACACAG-----AGGAACCGCTGCCAGCGG-----1810
 QY 135 ArgAlaGluAspProPheGlnGlyGluLeuLeuSerArgPheGlnIleSerGly 154
 DB 1811 -----CCGTCCCGCTTCAGCTGTGAGAGTGAAGTCTTCCAGAACTCCGCG 1858
 QY 155 TyrlleuArgTyrgSerCysGluSerArgIleArgSerTyrgLeuArgLysSerTyrg 174
 DB 1859 CATGTAAACCTGACCTTACAGATGATGATATAGAGAGAAAGAAATCGAGACTC 1918
 QY 175 ProSerThrValGlyAlaGluAlaGlnGluLeuArg-----188
 DB 1919 AAGAACGAGCTGACCTCTCTCCAGGCCGAGGGGAGTGAAGGACGAGAGCGGA 1978
 QY 189 -----ValleuGlySerMetCysGlnArgLeuArg 198
 DB 1979 GAGTCCCTGCGTAAAGCCGTACCTGCGGACGCTACAGGAAACGTGACTTGCACTTGA 2038
 QY 199 SerMetGlnTyrgAsnGlySerTyrgPheAspArgGlyAlaLysGlyGly-----214
 DB 2039 GCACTCAACAGGATGACAAATTCAAGCCCTCAGAGTTCAGCGGCAATATTCCTACT 2098
 QY 215 -----SerArgLeuCysThrProGlnGlyTyrgPheSerCysGlnGlyProPhe 230
 DB 2099 TGGAACTACGACCCCATGACTCTCC-----TCATCC-----2134
 QY 231 AspMetAspSerCysLeuSerArgHisSerIleAsnProTyrgSerAsnArgGluSerArg 250
 DB 2135 -----TCCTTTCCCGAGG-----TAC 2152
 QY 251 IleLeuPheSerThrTyrgAsnLeuAspHisIleIleGluLysTyrgThrIleIlePro 270
 DB 2153 CTCCTGTGACCTCTGCTGACTG-----CTCCCAATTAACTAGCGTTTCCCT 2203
 QY 271 ThrIleu 272
 DB 2204 TTCCTA 2209

RESULT 7
 US-08-630-524-1
 Sequence 1, Application US/08630524
 Patent No. 5690930
 GENERAL INFORMATION:
 APPLICANT: Chen, Jane J.
 APPLICANT: London, Irving M.
 TITLE OF INVENTION: DNA Encoding the Heme-Regulated
 Eukaryotic Initiation Factor 2 alpha kinase

NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Patrea L. Pabst
 STREET: 2800 One Atlantic Center
 STREET: 1201 West Peachtree Street
 CITY: Atlanta
 STATE: Georgia
 COUNTRY: US
 ZIP: 30309-3450
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/630,524
 FILING DATE: 10-APR-1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/938,782
 FILING DATE: 31-AUG-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Pabst, Patrea L.
 REGISTRATION NUMBER: 31,284
 REFERENCE/DOCKET NUMBER: W0101CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (404)873-8794
 TELEFAX: (404)873-8795
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2729 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FRAGMENT TYPE: N-terminal
 ORIGINAL SOURCE:
 CELL TYPE: Reticulocytes
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 113..2149
 OTHER INFORMATION: /note= "Expression of HRI
 OTHER INFORMATION: mRNA in Human erythroid cells, using as the
 OTHER INFORMATION: probe rabbit HRI cDNA from nucleotides 113
 OTHER INFORMATION: to 2149."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 229..249
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of human HRI cDNA sequence
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 543..560
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of human HRI cDNA sequence
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 448..468
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of human HRI cDNA sequence
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1009..1031
 OTHER INFORMATION: /note= "Primer used in the
 OTHER INFORMATION: amplification of a human HRI cDNA sequence
 OTHER INFORMATION: using the rabbit HRI cDNA sequence."
 PUBLICATION INFORMATION:

AUTHORS: Chen, Jane J.
 AUTHORS: London, Irving M.
 TITLE: Cloning of the cDNA of the heme-
 TITLE: regulated eukaryotic initiation factor
 TITLE: 2alpha (eif-2alpha) kinase of rabbit reticulocytes:
 TITLE: Homology to yeast GCN2 protein kinase and human
 TITLE: double-stranded-RNA-dependent
 JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
 VOLUME: 88
 PAGES: 7729-7733
 DATE: September-1991
 US-08-630-524-1
 Alignment Scores:
 Pred. No.: 0.247
 Score: 96.00
 Percent Similarity: 35.438
 Best Local Similarity: 20.538
 Query Match: 5.378
 DB: 1
 Gaps: 13
 US-09-748-451-2 (1-338) x US-08-630-524-1 (1-2729)
 QY 4 LysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyAlaGly 23
 DB 1427 AACCCAGAAATATTTCTTCATGCTCATACAGTGAAGTGAATAGAGACTTGTG 1486
 QY 24 ArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPhe-----Gln 38
 DB 1487 CTGGCCCTGCGCCGACATCCAGAAATGCGCGGACGACGAGAAACGGGGAGAGA 1546
 QY 39 LeuProGluArgLysSerArgLeu-----CysLeuArgGluAspGlyThrGluLeuThr 56
 DB 1547 GCACCCACACACACTTCGAGTGGGACACTGTGTATACCCCTGCCCGGACGAGTTGGA 1606
 QY 57 GluAspThrPheProSerValProAsp-----AsnAlaGluLeuValLeuLeuThr 74
 DB 1607 GCATCGAGATATATGCCAAGTCACATGACACCTCGCGCTGATCTCTGACACTC 1666
 QY 75 GlyAlaIleArgPheGlnGlyTyrValSerAspIleArgPheLeuSerAlaPheIleGlu 94
 DB 1667 -----TTCACGCCCTTCGCGACACAGATGAGCGG---CGAAGGCTCTGACGGGC 1714
 QY 95 ProGlnValGlyLeuIleGlnAlaAlaGlnLeuLeuGlyAspGlnAlaProGln 114
 DB 1715 GTGGACCTGCGCGCATACCGCATCTCCTCAGTAAAGAGTGGCGGCGGACCAAGTAC 1774
 QY 115 ArgGlnArgLeuLeuAlaAspLeuIleAsnValSerGlnAsnIleAlaAlaGluThr 134
 DB 1775 GTCCACGCTGCTGACACGAG-----AGCAACCGCTCCACGCG----- 1810
 QY 135 ArgAlaGluAspProPheGlnGlyLeuGlnGlyLeuSerArgPheGlnSerLysGly 154
 DB 1811 -----CCGTCCGCGCTTCAGCTGCTGAGAGTGTGAGTCTTCAGAAACCTCCGC 1858
 QY 155 TyrLeuArgTyrSerCysGlnSerArgTyrIleArgSerTyrLeuArgGluValSerSerTyr 174
 DB 1859 CATTTTAACTCACTACCTTACAGATGAGTATTAAGACGAAAGAAATATCAGACACTC 1918
 QY 175 ProSerThrValGlyAlaGluAlaGlnGluIleLeuArg----- 188
 DB 1919 AAGAAAGACGCTGAGCTCTCTCCAGGCGCGGAGGGGTGAGGTGACAGCGAGACGGA 1978
 QY 189 -----ValLeuGlySerMetCysGlnArgLeuArg 198
 DB 1979 GAGTCCCTGCTGACGCTGACCTGCGACGACGTCACAGGAGACCTGACACTGCA 2038
 QY 199 SerMetGlnTyrAsnGlySerTyrPheAspArgGlyAlaAlaGlyGly----- 214
 DB 2039 GCAGTCAACTGAGATGAGCAATTTCAACCTCTGAGATTCAGCGGCAATATCTCTACT 2098
 QY 215 -----SerArgLeuGlyThrProGluGlyTyrPheSerCysGlnGlyProPhe 230

Db 2099 TGCATGACTCAGCCCGCATGACTCTCCCC-----TCATGC----- 2134
QY 231 AsphetaSerCysLeuSerArgHisSerIleAsnProTyrSerAsnArgGluSerArg 250
Db 2135 -----TCTCTCTCCCGGAGG-----TAC 2152
QY 251 IleLeuPheSerThrTrpAsnLeuAspHisIleIleGluLysAspGlyIleIlePro 270
Db 2153 CTCTGTGTGACCTCTGTGTGACTG-----CTCCCAATTAACTTACGCTTTCCCT 2203
QY 271 ThrLeu 272
Db 2204 TTCCTA 2209
RESULT 8
PCT-US93-08131-1
Sequence 1, Application PC/TUS9308131
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute Of Technology
TITLE OF INVENTION: DNA Encoding the Heme-Regulated
NUMBER OF SEQUENCES: 2 Eukaryotic Initiation Factor 2 alpha Kinase
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kilpatrick & Cody
STREET: 1100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: US
ZIP: 30309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08131
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MIT 5807
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-815-6508
TELEFAX: 404-815-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2729 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Rabbit
CELL TYPE: Reticulocytes
FEATURE:
NAME/KEY: misc.feature
LOCATION: 113..2149
OTHER INFORMATION: /note= "Expression of HRI mRNA in Human erythroid cells,
FEATURE:
NAME/KEY: misc.feature
LOCATION: 229..249
OTHER INFORMATION: /note= "primer used in the amplification of human HRI cdn
FEATURE:
NAME/KEY: misc.feature
LOCATION: 543..560
OTHER INFORMATION: /note= "Primer used in the amplification of human HRI cdn
FEATURE:
NAME/KEY: misc.feature
LOCATION: 448..468

OTHER INFORMATION: /note= "Primer used in the amplification of human HRI cdn
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1009..1031
OTHER INFORMATION: /note= "Primer used in the amplification of a human HRI c
PUBLICATION INFORMATION:
AUTHORS: Chen, Jane J.
TITLES: Cloning of the cDNA of the heme-regulated
TITLES: Eukaryotic Initiation factor 2alpha
TITLES: (eif-2alpha)kinase of rabbit reticulocytes:
TITLES: Homology to yeast GCN2 protein kinase and human
JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
VOLUME: 88
PAGES: 7729-7733
DATE: September-1991
PCT-US93-08131-1
Alignment Scores:
Pred. No.: 0.247 Length: 2729
Score: 96.00 Matches: 62
Percent Similarity: 35.43% Conservative: 45
Best Local Similarity: 20.53% Mismatches: 121
Query Match: 5.37% Indels: 74
Gaps: 13
US-09-748-451-2 (1-338) x PCT-US93-08131-1 (1-2729)
QY 4 LysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyValAlaGly 23
Db 1427 AAGCTAGAAATATTTTCTTCATGCTGATCAACAAAGTAAATAGGACTTGGT 1486
QY 24 ArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPhe 38
Db 1487 CTGGCTTCGCGGACATATTCACAGAAATCGGCGCCGACACAGAAAGGAGAGA 1546
QY 39 LeuProGluArgGlySerArgLeu-----CysLeuTyrGluAspGlyThrGluLeuThr 56
Db 1547 GCACCCACACACACTTCCCGAGTGCGACCTGTGTACGCGCTCGCCGACAGTGGAA 1606
QY 57 GluAspTyrPheProSerValProAsp-----AsnAlaGluValLeuLeuThr 74
Db 1607 GGATCGAGTATGATGCCAGTACAGTACAGCGTGGCGGTGATCTCTGAGACTC 1666
QY 75 GlyAlaIleArgGlnGlyValSerAspIleArgPheLeuSerAlaPheIleGlu 94
Db 1667 -----TTCACGCCCTTGGGACAGAGTGGAGCG-----GCAGAGTCTCTGAGGCG 1714
QY 95 ProGlnValGlyLeuIleGlnAlaAlaGlnIleLeuLeuCysAspGluGlnAlaProGln 114
Db 1715 GTGGAGCTGGCCGATCCGACTCCCTCAGTAAGAGTGCAGCGGCGCAGGCAATGAC 1774
QY 115 ArgGlnArgLeuLeuAlaAspLeuLeuHisValSerGlnAsnIleAlaIleGluThr 134
Db 1775 GTCCAGCTGTGACCAAG-----AGGAACCGCTCCGCG----- 1810
QY 135 ArgAlaGluAspProTyrPheGlnGlyLeuGlnLeuSerArgPheGlnIleSerLys 154
Db 1811 -----CCGTCCGCCCTTCAGCTGTCGACAGTGCCTTCCAGAACTCCGCG 1858
QY 155 TyrLeuArgTyrSerCysGluSerArgIleArgSerTyrLeuArgGluValSer 174
Db 1859 CATGTAACTACACCTACAGATGATATATAGCAGCAAGAAAGAAATGAGGAATC 1918
QY 175 ProSerThrValGlyAlaGluAlaGlnGluLeuLeuArg----- 188
Db 1919 AAGAAGCAGTACGCTCTCTCCAGGCCGAGAGGAGAGTACAGGAGGAGACGCA 1978
QY 189 -----ValLeuGlySerMetCysGlnArgLeuArg 198
Db 1979 GAGCTCCCTGCTACGCTACCTGCGGCGACGTCACAGGGGAAAGCTGACCTTGCA 2038

```

QY 199 SerMetGlnTYrAsnGlySerTYrPheAspArgGlyAlaLysGlyLys----- 214
Db 2039 GCAGTCACTGGAATGAGCAATTTCAAGCCTCCTGAGTTCAGGGGAGATATCCTCACT 2098
QY 215 -----SerArgLysCysThrProGluGlyTYrPheSerCysGlnGlyProPhe 230
Db 2099 TGGAAATCACTGAGCCCGCATGACTCTCCCTC-----TATATG----- 2134
QY 231 AspMetAspSerCysLysSerArgHisSerIleAsnProTYrSerAsnArgGluSerArg 250
Db 2135 -----TGCCTCTCCCGAGG-----TAC 2152
QY 251 IleAspHisSerThrTrpAsnLeuAspHisIleIleGluLysArgThrIleIlePro 270
Db 2153 CTCTGAGTACCTCTGCTGACTG-----CTCCCAATTAACTTACCTTTTCCCT 2203
271 ThrLeu 272
2204 TTCCTA 2209

RESULT 9
US-09-069-023-21
Sequence 21, Application US/09069023A
Patent No. 6348573
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
APPLICANT: Koseki, Takeyoshi
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
FILE REFERENCE: UM-03333
CURRENT APPLICATION NUMBER: US/09/069,023A
CURRENT FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 2143
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-069-023-21

Alignment Scores:
Pred. NO.: 0.293 Length: 2143
Score: 94.00 Matches: 66
Percent Similarity: 36.25% Conservative: 50
Best Local Similarity: 20.62% Mismatches: 113
Query Match: 5.25% Indels: 94
Gaps: 12
9-748-451-2 (1-338) x US-09-069-023-21 (1-2143)
QY 6 LysSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyValAlaGlyArgSer 25
Db 269 AAGCCCTTCAAGTCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 328
QY 26 CysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuProGluArgGlySerArg 45
Db 329 CTGGAGAGATCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
QY 46 LeuCysLeuTYrGluAspGlyThrGluLeu-----ThrGluAspTYrPheProSerValPro 64
Db 389 ATCCACCTGAGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
QY 65 AspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTrpGlnGlyTYrValSerAsp 84
Db 449 GAGAAATAGGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 490
QY 85 IleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAlaAlaGln 104
Db 490 ----- 490
QY 105 GlnLeuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuLeuAlaAspLeuLeuHis 124
111

```

```

Db 491 -----ATGATCCCATCTAC 505
QY 125 AsnValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspProProTYrPheGluGly 144
Db 506 TACGTACAGAT-AACCACTCCACATGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
QY 145 LeuGlnSerArgPheGlnSerLysSerGlyTYrLeuArgTYrSerCysGlnSerArgIle 164
Db 559 GCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 165 ArgSerTYrLeuArgGluValSerSerTYrProSerThrValGlnAlaGlnAlaGlnGlu 184
Db 601 -----CTCGAGATCGCTCGATCCCGCATCGCTGAGAGAGAGAGAGAGAGAGAGAG 648
QY 185 GluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTYrAsn--- 203
Db 649 CCTCTGCAAGGTCGTCGATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGTCC 708
QY 204 -----GlySerTYrPheAspArgGlyAlaLysGlyLysArg----- 216
Db 709 CAACCTGTTGTCGATACACGAGAAAGAGTTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 768
QY 217 -----LeuCysThrProGluGlyTYrPheSerCysGlnGlyProPheAspMetAspSer 234
Db 769 GCCATTGTGTGCCAAGCGGAA-----TCCGAGAGATCGTTAAATCTGCTGAA 816
QY 235 CysLeuSerArgHisSerIleAsnProTYrSerAsnArgGluSerArgIleLeuPheSer 254
Db 817 GCTGA-AGCTATTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 848
QY 255 ThrTrpAsnLeuAspHisIleIleGluLysArgThrIleIleProThrLeuValGlu 274
Db 849 -----GCTCGAAGAGTTATCCCGAGAGTGG- 877
QY 275 AlaIleLysGluGlnAspGlyArgGluValAspTrpGluTYrPheTYrGlyLeuLeu--- 293
Db 878 GCCATTGACCGGAGAGATGAGAGAGAGTGAAC---GAGCGGAGAGAGAGAGAGAGAGAG 934
QY 294 ---PheThrSerGlnAsnLeuLysLeuValHisIleVal---CysHisLysLysThrThr 311
Db 935 ACCCAACAGCAGCAGCAACCCGCAACATCAGTGTGATGAGACATCAACCAACAGC 994

RESULT 10
US-09-069-023-19
Sequence 19, Application US/09069023A
Patent No. 6348573
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
APPLICANT: Koseki, Takeyoshi
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
FILE REFERENCE: UM-03333
CURRENT APPLICATION NUMBER: US/09/069,023A
CURRENT FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 1167
TYPE: DNA
ORGANISM: Mus musculus
US-09-069-023-19

Alignment Scores:
Pred. NO.: 0.256 Length: 1167
Score: 91.00 Matches: 47
Percent Similarity: 40.61% Conservative: 20
Best Local Similarity: 28.48% Mismatches: 63
Query Match: 5.09% Indels: 35
Gaps: 8
US-09-748-451-2 (1-338) x US-09-069-023-19 (1-1167)

```



```

Db      1082 -----CTTGAGAGACCCACCCTCATCGCTGTGCAGAGATCCACAGATCATCTTCTC 1132
Oy      263 GtUyLySArGThIleIleProThIleValGluAlaIleLysGluGlnAspLyArg 282
Db      1133 TCCAAAGTCAGCTGCTCCGTTCCAGCTGCTCGAGCACTG-----CAAAAGAACGGACAGC 1186
Oy      283 GlUValAspTrpGluTyRPhetYglYleuLeuPheThSerClnuSlnLeuYsLeuVal 302
Db      1187 CAG----- 1189
Oy      303 HsILleValCYsHsILys-----LysThrThrHisILysIleuAsn-CysAspPro-- 318
Db      1190 TATCTGGCCTGCCCAAGAGATATATAGAGACTTCTGCATCCATCCACTGCTTTCGCCAC 1249
Oy      319 -----SerAylIleTyLysProGlnThThArgLeuYsArgLyGcl 332
Db      1250 TTGACCGCGTGACACCTGTATCATCAGAGGAGTGCACACGCTCGGGCCCCACACGAGCA 1309
Oy      332 nProValArgLyAsArgIln 338
Db      1310 GCCGACTCGTTCAGCGAGG 1328

RESULT 12
US-09-061-702-3
; Sequence 3, Application US/09061702
; Patent No. 6165737
; GENERAL INFORMATION:
; APPLICANT: Wang, Xiaodong
; APPLICANT: Liu, Xuesong
; TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210-4433
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/061,702
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: McMillian, Nabeela R.
; REGISTRATION NUMBER: P-43,363
; REFERENCE/DOCKET NUMBER: UTSD:546
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512)418-3000
; TELEFAX: (512)474-7577
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1689 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 113..1105
; US-09-061-702-3

Alignment Scores:
Pred. No.: 0 814
Score: 89.00
Percent Similarity: 36.18
Best Local Similarity: 22.228
Query Match: 4.978

Length: 1689
Matches: 78
Conservative: 49
Mismatch: 146
Indels: 78

```

Db 977 TGTAGCGGAGCTGCCCTGCGCTGACAG 1009

RESULT 13

5223423-1

Patent No. 5223423

APPLICANT: FRANCHINI, GENOVEPPA, MONG-STAL, FLOSSIE;

GALLO, ROBERT

TITLE OF INVENTION: CHARACTERIZATION OF REPLICATION COMPETENT HUMAN IMMUNODEFICIENCY TYPE 2 PROVIRAL CLONE HIV-2 SBL/ISY

NUMBER OF SEQUENCES: 4

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/331,212

FILING DATE: 03-31-1989

SEQ ID NO: 1

LENGTH: 9633

Alignment Scores:

No.	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
15.6	9633	40	20	55	29	7
88.50						
41.67%						
Local Similarity:						
4.95%						

Query Match: 6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

6

US-09-748-451-2 (1-338) x 5223423-1 (1-9633)

16 Proctarglyphegylvalalaglyargsercysglnluvalleuarglyglsleu 35

2658 CCACACCCAGCAGGATTAAGCAGAGAAAGAAATTAAGTGTAGAGATGCC 2717

36 Argpheglnleuprogliuarglyserarglyleuarglyglsleu 55

2718 TACTTTCCATACCC-----CTATATGAGATTTTACACAGAT 2756

56 Thrgluapryrpherpservalproaspasnalaglu-----leuvalleu 71

2157 ACTGATTTACTGTCATCAGTA--ACATGACAGAGAGAAAGATATATAC 2813

72 Leuthrleuglinalatrpnglnly-----tyrvalseraspillearg 87

2814 AAGGTCTTACACAGGATGAGAGGCTCACACCATTTTCAATACATAGGCA 2873

88 Pheleuseralphehlsu-----proglvalleuileglnalalagln 105

2874 GCTTAGAACATTCAGAAAGAACACCATGTCATTTAGTACATGATGAT 2933

106 Leuleucysapgluglnalarglnarglyleu-----leuvalleu 118

2934 ATCTTAATA-----GCTAGTACAGACAGATTTGACAAATGACAAAGTCTTA 2984

119 Leuvalaspleulehlsasvalsercylasnllealalagluthrargalaglu 137

2985 CAGGTAAAGAACTTCATTAATGGCTGGATTTTCCACCCAGAGAAAGTCCAAAG 3044

138 Aspproptp 141

3045 GACCTCCATAC 3056

RESULT 14

US-09-069-023-13

Sequence 13, Application US/09069023A

Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohiro

APPLICANT: Koseki, Takeyoshi

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

FILE REFERENCE: UM-03333

CURRENT APPLICATION NUMBER: US/09/069,023A

NUMBER OF SEQUENCES: 3

SEQUENCE 13, Application US/09069023A

Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel

APPLICANT: Inohara, Naohiro

APPLICANT: Koseki, Takeyoshi

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS

FILE REFERENCE: UM-03333

NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 1049
TYPE: DNA
ORGANISM: Homo sapiens
US-09-069-023-13

Alignment Scores:

Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
0.5	1049	61	31	85	37	10
88.00						
43.19%						
Best Local Similarity:						
28.64%						
Query Match:						
4						

US-09-748-451-2 (1-338) x US-09-069-023-13 (1-1049)

14 ArgserProarglyphegylvalalaglyargsercysglnluvalleuarglyglsleu 33

195 AGGACAGAGCCGCGCT---GGGCTGATGAGCAGAGCCTGAGAGATCATCAGCAAGCT 251

34 Cysleuargpheglnleuprogliuarglyserarglyleuarglyglsleu 53

252 CTGATGCCCTGCTATCGCTACCGGACCTGCTACTGCTGCTGAGAGAGATGACAC 311

54 Gluleu---Thrgluapryrpherpservalproaspasnalaglu-----leuvalleu 72

312 GTGCTGACACAGAGAGATTTTACAGACCTTGGAGACCAACAGCATTCATCATCTTG 371

73 Thrlleuglinalatrpnglnlytyrvalseraspilleargargpheleuseralphe 92

372 GAAAGAGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 430

93 His-----gluproglnvalleuileglnalalaglnleuvalleu 108

431 GAGTGGGATATACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 487

109 Aspluglnalalarglnalarglnalarglnalarglnalarglnalarglnalargln 128

488 ---CATGCGCTTAACTGAG 541

128 nasnllealalagluthrargalagluaspproptp 148

542 CGACATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 571

148 gphleuserlysergyltyrleuargtyrsercysgluserarglyleuargtyrleu 168

572 GCTAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 628

168 uargluvalsergyltyrleuargtyrsercysgluserarglyleuargtyrleu 188

629 CATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655

188 gvalleuglysermetcysglnarg-----leuargsermetcyltyrserg 206

656 GGTCTGATGACAG 715

206 yrphesaparglyalalaglyserarglyleuargtyrleuargtyrleu 217

716 GTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747

RESULT 15

US-08-794-158-1

Sequence 1, Application US/08794158

Patent No. 6387655

GENERAL INFORMATION:

APPLICANT: Burnett Jr., J. Paul

APPLICANT: Mayne, Nancy G.

APPLICANT: Sharp, Robert L.

TITLE OF INVENTION: Excitatory Amino Acid Receptor Protein

SEQUENCE 1, Application US/08794158

Patent No. 6387655

GENERAL INFORMATION:

APPLICANT: Burnett Jr., J. Paul

APPLICANT: Mayne, Nancy G.

APPLICANT: Sharp, Robert L.

TITLE OF INVENTION: Excitatory Amino Acid Receptor Protein

SEQUENCE 1, Application US/08794158

Patent No. 6387655

GENERAL INFORMATION:

APPLICANT: Burnett Jr., J. Paul


```

OY 25 SerCysGlnValLeuArgLysGlyCysLeuArgPheGlnLeuProGlnArgLysSer 44
DB 1421 AAAAAATTGAAGTAAAGATTGGATGTTAGCAAGAAATTC---CATTCAGATACG 1477
OY 45 ArgLeuCysLeuTyrGluAspGlyThrGlu-----LeuThrGluAspTyrPheProSer 62
DB 1478 TCGCATTCCTCAAGAAATCAACAGAAAGATCTTCACAGAAAGATATATATGCGCAAT 1537
OY 63 ValProAspAsnAlaGlu---LeuValLeuLeuThrLeuGlyGlnAlaIleProGlnTyr 81
DB 1538 GGTAAAGATGATGAAACAAAGAAATATATATGTTAGTGTATGATCAAAACAAAGA----- 1591
OY 82 ValSerAspIleArgPheLeuSerAlaPheHisGlu----- 94
DB 1592 -----AAAATTCCAAACATCTTACATGACCAAGAAAGAAAGATGGGTAC 1639
OY 95 ProGlnValGlyLeu-----IleGlnAlaAlaGlnGln-----LeuLeuCys 108
DB 1640 CTTAAATATAGTCTGCGCAGGATGGATGCGATGACAGAGATGCTCATTCACCAAGAC 1639
OY 109 AspGlnGlnAlaProGlnArgGlnArgLeuLeuAlaAspLeuLeuHisAsnValSerGln 128
DB 1700 GAACATACCGCATCCCAACCCGGGCATCAACAGATGCTCATTCACCAAGAC 1759
OY 129 AsnIleAlaIleGluThrArgAlaGluAspProTyrPheGlnGlyLeuGlnSerArg 148
DB 1760 AATTAAAGTTGGAA---GAAGCTGTGTAAACCCAGACATCATTCATTCAGTGA 1816
OY 149 PheGlnSerLysSerGlyTyrLeuArgTyrSerCysGlnSerArgIle----- 164
DB 1817 GCGACGCTTATGAGGAGATACCTGTCATAGTTGTTATCAAGAACTGAAATGTTTG 1876
OY 165 ---ArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAlaGlnAla 183
DB 1877 AGCAATGATTCCTTTCCTCAAGTTCAGAAATAAATCAAGTGGGATTAATTCGAAATCAG 1936
OY 184 GlnGlnPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetLeuIlePyrAsn 203
DB 1937 GAT----- 216
OY 204 GlySerTyrPheAspArgGlyAlaLysGlyGlySerArg----- 216
DB 1958 CAAAGTTGGAATGTTCTGGAAGATGACAGGAAATCAAGAGAAATCAAGAACTATAC 2017
OY 216 ----- 216
DB 2018 GACATATCAAGCAAAACAGATTCGAAAGCTGGGGAAGAAAGATATATCCAGG 2077
OY 217 -----LeuCysThrProGlnGlyTyrPheSerCysGlnGlyProPheAspMetAsp--- 233
DB 2078 AGTTGAAAGCTACTACAGAAATGTTTACCTGCCAAGAAACAGTGTGTGTAAGT 2137
OY 234 SerCysLeuSerArgHisSerIle----- 241
DB 2138 TCTTCTCTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2197
OY 242 -----AsnProTyrSerAsnArgGlnSerArgIleLeuPhe 253
DB 2198 AAGACAACATCAGAAAGTACTCCAGAAAGCATGCTGCTGACAAAGAAACATTAATCT 2257
OY 254 -----SerThrTyrAsnLeuAspHisIleIleGlnLysLysArgThrIleIle 269
DB 2258 AAGCTACTCAGAGACAGCAAGAGTACAGGCGCCATCGAGTAAAGAAACA----- 2311
OY 270 ProThrLeuValGlnAlaIleLysGlnGlnAspGlyArgGlnValAspIlePyrGly--- 288
DB 2312 -----GCGTTGATCCACATGAAAGGGAAGAAATGATCATCATTCATCACC 2356
OY 289 -----PhenylGlyLeuLeuPheHisSerGlnLysLeuValHis 303
DB 2357 CTTTGCAACGATACAGTACGTAGTGTATCTATGACAAATATTTTGAAGAAATCACCT 2416

```

```

OY 304 Ile---ValCysHisLysLysThrHisLysLeuAsnCysAspProSerArgIleTyr 322
DB 2417 TTACGATATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2476
OY 323 LysProGlnThrArgLeuArgLysGln 332
DB 2477 AATCTTACGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2506

```

RESULT 18
 US-09-069-023-17
 ; Sequence 17, Application US/09069023A
 ; Patent No. 6348573
 ; GENERAL INFORMATION:
 ; APPLICANT: Nunez, Gabriel
 ; APPLICANT: Inohara, Naohiro
 ; APPLICANT: Koseki, Takeyoshi
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
 ; FILE REFERENCE: US-03333
 ; CURRENT APPLICATION NUMBER: US/09/069,023A
 ; CURRENT FILING DATE: 1998-04-27
 ; NUMBER OF SEQ ID NOS: 38
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 17
 ; LENGTH: 1114
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-069-023-17

Alignment Scores:
 Pred. No.: 0.967 Length: 1114
 Percent: 86.00 Matches: 65
 Percent Similarity: 37.55 Conservative: 27
 Best Local Similarity: 26.53 Mismatches: 102
 Query Match: 4.81% Indels: 52
 DB: 4 Gaps: 13

US-09-748-451-2 (1-338) x US-09-069-023-17 (1-1114)
 OY 14 ArgSerProArgLysPheGlyValAlaIleArgSerCysGlnGlnValLeuArgLysGly 33
 DB 343 CGAAGTACCGCGCGT---GGGTGATGGCCACAGCTGTCGAGAACTTATACGAAAGCT 399
 OY 34 CysLeuArgPheGlnLeuProGlnArgGlySerArgLeuCysLeuTyrGlnAspGlyThr 53
 DB 400 CTGATGCTTACTGCTCAACAGTGGCTGTTACGCTGCTGCTGAGAGAGAGCGGAC 459
 OY 54 GluLeu---ThrLysPyrPheProSerValProAspAsnAlaGlnLeuValLeuLeu 72
 DB 460 GTGCTGACACAGAGAGATTTCTGACACTTAAGGACAAACAGCATTTCTATGATCTTG 519
 OY 73 ThrLeuGlyGlnAlaIleP-----GlnGlyTyr 81
 DB 520 GAAGAGGACGAAATGACACACCGGCTAGTATGATGATGATGATGATGATGATGATGAT 579
 OY 82 ValSerAspIleArgArg-----PheLeuSerAlaPheHisGlnPyrProGlnValGlyLeu 99
 DB 580 AATCGGGAATCCAGAGTACACTTCA---CCTTACAGCTGAAACCCCA---GGACTT 635
 OY 100 IleGlnAlaIleGlnGlnLeuLeuLeuCysAspGlnAlaIlePyrProGlnArgIleLeu 119
 DB 636 CCTCGGCTGCTCA-----TGT-----CAAGCCAGCATGTACAGATATATC 680
 OY 120 AlaAspLeuLeuHisAsnValSerGln-----AsnIleAlaIleGlnThr 134
 DB 661 GGTCTCTACGACATCCATGACACAGGTTCAAGCCGCTGTAAAGAAATCTCTAGGCTT 740
 OY 135 Arg-----AlaLysAspProTyrPheGlnGlyLeuGlnSerArgPheGlnSer 151
 DB 741 TATCTCTATGCTGACAGATGATGACGAGCACTTCTGCTGCTA----- 782
 OY 152 LysSerGlyTyrLeuArgTyrSerCysGlnSerArgIleArgSerTyrLeuArgGlnVal 171

Db 783 ---TGGGGGACATACATGCTCCGAGTACTGGCGATACAGAGA----- 824
Oy 172 SerSerTyProSerThValGlyAlaGlnAglngluGluPheLeuArgValLeuGly 191
Db 825 -----CGAGCATATCCCAAGACCTAGACCAAGAGCGTGTATCATGTACAGG 872
Oy 192 SerMetCysGlnArgLeuArgSerMetGlnTyTrasnGlySerTyTrpPheAsp---ArgGly 210
Db 873 GCACAGCTACAGAGCGCCAGGACCTGCTCTGTATTATAGCTGTGGATGCCAGGGA 932
Oy 211 AlalysGlyGlySerArgLeuArgCysThrProGluGlyTrpPheSerCys----- 226
Db 933 AGGATGGGGGTGTATCCCATGTGAGGCTGATAGACAGATTCCTGCAAGAAAGGCG 992
Oy 227 ---GlnGlyTrpPhe 230
Db 993 GCAGAGGGGCGCTTTC 1007
RESULT 19
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
Alignment Scores:
Pred. No.: 4.86e+05 Length: 4403765
Score: 85.50 Matches: 67
Percent Similarity: 36.64% Conservative: 29
Best Local Similarity: 25.57% Mismatches: 89
Query Match: 4.78% Indels: 77
Gaps: 15
-748-451-2 (1-338) x US-09-103-840A-2 (1-4403765)
2 LeuGlnLysProLysSerValLysLeuArgAla-----LeuArgSerProArgLys 18
Db 201968 CTGAATCTGCCAAGATGCTCAGGGTTTCAGCCAAATCTCTCTGCGGTACCAACAGG 201909
Oy 19 PheGlyValAlaGlnArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGln 38
Db 201908 ---TCGGTGAAGAGTCCAGCTGCTCGAGCACA-----TGGTTCAGGTTTCGGG 201864
Oy 39 LeuProGluArgGlySerArgLeuArgCysLeuArgLysGlnGluValLeuArgLysGln 58
Db 201863 TTGGTCGTTG-----ATTAGGTTTTCACCTGGGCGCAGAAAGCGAGATG----- 201819
Oy 59 TyPheProSerValProAsp-----AsnAlaGluLeuValLeu 71
Db 201818 ---TTCCCAAGAGGGGCTGACCGCGCGCGGCTGTGTAACGCGCGCATAGGCTC 201762
Oy 72 LeuThrLeuGlnAlaIATrGlnGlyTyTrValSerAspIleArgArgPheLeuSerAla 91
Db 201761 TTAGCGTTGACCAT-----AGCGGTGACCTGCTCA 201729

Oy 92 PheHisGluProGlnVal-----GlyLeuIleGlnAlaAlaGlnGlnLeuLeuCys 108
Db 201728 CTGGATACACCCAGAGATCTGCGCACCTGTTGGCTGGCGGTAGTACGTATGATCTGC 201669
Oy 109 AspGluGlnAlaProGlnArgGlnArgLeuLeuAlaAspLeuLeuHisAsnValSerGln 128
Db 201668 TCGTCGCGCTTGCAGT-----GTCTCGAG 201642
Oy 129 AsnIleAlaAlaGluThrArgAla-----Glu 137
Db 201641 AACTTAGCACCCTGTCAGGCGCGCTCAGTATGCTGATCAACGCTCC 201582
Oy 138 AspProTrpTrpPheGluGlyLeu-GluSerArgPheGlnSerLysSerGlyTyTrpLeuAr 157
Db 201581 GACAACATCATCAGCAGACCTGACCGCTCA----- 201548
Oy 157 gTySerCysGlnSerArgIleArgSerTyTrpLeuArgLysValSerSerTyTrpProSerTh 177
Db 201547 ---TGTCCACCGGATGCGGCTTGTGATGACGTGAGAAAGAGGCTGTAATCTGG 201495
Oy 177 rValGly-----AlaGluAlaGlnGluPheLeuArgValLeuGlySerMet 193
Db 201494 TAGGGGTGTGCTTGTGCCAAGCGCAAGCGCCCGGCGGCAACGCTTGGCGCGC 201435
Oy 193 tCysGlnArgLeuArgSerMetGlnTyTrasnGlySerTyTrpPheAspArgGlyAlaLysGln 213
Db 201434 CGCGCTGCAATCTGAGCAGCTTCTCTACCCAGGA---TGGTGTGCTGCGGATGGCTAGG 201378
Oy 213 yGlySerArgLeuArgSerThrProGluGlyTrpPheSerCysGln-GlyProPheAspMetA 233
Db 201377 CGGCTCTCGG---TGCCGA-----TGGTGTGCTGCGGATGGAGAACTTACACAGC 201330
Oy 233 sp 233
Db 201329 AT 201328
RESULT 20
US-09-103-840A-1/c
Sequence 1, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/103.840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 4411529
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 4.87e+05 Length: 4411529
Score: 85.50 Matches: 67
Percent Similarity: 36.64% Conservative: 29
Best Local Similarity: 25.57% Mismatches: 89
Query Match: 4.78% Indels: 77
Gaps: 15
US-09-748-451-2 (1-338) x US-09-103-840A-1 (1-4411529)
Oy 2 LeuGlnLysProLysSerValLysLeuArgAla-----LeuArgSerProArgLys 18
Db 201801 CTGAATCTGCCAAGATGCTCAGGGTTTCAGCCAAATCTCTCTGCGGTACCAACAGG 201742

```
OY 19 PheGlyValAlaIleArgSerCysGlnIleValLeuArgLysGlyCysLeuArgPheGln 38
DB 201741 ---TCGGTAGAGATCGCGAGCTGCTCAGACACA-----TGGTTCAGGTTCCGG 201697
OY 39 LeuProGlnIleArgGlySerArgLeuCysLeuTyrGlnAspGlyThrGlnLeuThrGlnAsp 58
DB 201696 TTGTCGTTG-----ATTAAGGTTTGTGACCTGGGCGGAGAAAGCGAGATG----- 201652
OY 59 TyrPheProSerValProAsp-----AsnAlaGlnLeuValLeu 71
DB 201651 ---TTCCCAAGAGGGGCGTCACCGCGCGCGCTGCTGTAACCGCGCATCAGGCTC 201595
OY 72 LeuThrLeuGlyGlnAlaIleTyrGlnGlyTyrValSerAspIleArgArgPheLeuSerAla 91
DB 201594 TTAGCGTTGACCAAT-----AGCGCGTCGACCTGCTCA 201562
OY 92 PheHISGlnProGlnVal-----GlyLeuIleGlnAlaIleGlnLeuLeuCys 108
DB 201561 CTCGATACACCGAGATGCTGCGCACCTGGTGGCTAGTAGGTCGATCTGC 201502
OY 109 AspGlnGlnAlaProGlnIleArgGlnArgLeuLeuAlaAspLeuLeuHISAsnValSerGln 128
DB 201501 TCGTCGCGCTTCCGATG-----GTGTCGGAG 201475
OY 129 AsnIleAlaIleGlnThrArgAla-----Glu 137
DB 201474 AACTTAGCCACCCCGTCGAGGGCGCGCTCAGTGCAGTAGGTCGATCAACGGCTCC 201415
OY 138 AspProProTyrPheGlnGlyLeu-GluSerArgPheGlnSerLysSerGlyTyrLeuArg 157
DB 201414 GACACACATTCAGCGACCGCTGACCTCGA----- 201381
OY 157 GlySerCysGlnSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThr 177
DB 201380 -----TGTCCACACCGGATGCGCCTTGTGTACGTGCAGAGACCGCTGTAAATCTGG 201328
OY 177 rValGly-----AlaGlnAlaGlnGlnGlnIleValLeuArgValLeuGlySerMe 193
DB 201327 TACGGGGTGTGCTTTGCCAACCGCAAAACCGCCCGCGGCAACGCTGGGCGCCCG 201268
OY 193 tCysGlnArgLeuArgSerMetGlnTyrAsnGlySerTyrPheAspArgGlyAlaLysGln 213
DB 201267 CGCGGCTCGATCTCGAGACATTCTCTCCACAGGA--TGGTGTGCGGTGCGGATGGCTAGG 201211
OY 213 xGlySerArgLeuCysThrProGlnGlyTyrPheSerCysGln-GlyProPheAspMetA 233
DB 201210 CGGCTCTCGG--TCCGA-----TGGTGTGTGTCGATGAGACACTGACACAG 201163
OY 233 sp 233
DB 201162 AT 201161
```

Search completed: May 26, 2003, 17:28:03
Job time : 2100 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 15:26:43 ; Search time 197 Seconds
(without alignments)
2265.565 Million cell updates/sec

Title: US-09-748-451-2

Perfect score: 1789

Sequence: 1

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODL-frame_plus_p2n.model -DEV-xlh

-O-cgna2.1/USPTO.spool/US09748451/runat.21052003.153830.16347/app.query.fasta.1.519

-DB-Published.Applications.NA -OFM-fastlap -SUPPLX-ropb -MINMATCH-0.1

-LOOCC-0 -LOOEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62

-TRANS-human40.cst -LIST-45 -DOCALLIGN-200 -THR.SCORE-pct -THR.MAX-100

-THR.MIN-0 -ALIGN-20 -MODE-LOCAL -OUTPMT-pto -NORW-ext -HEA5TIF-500 -MINLEN-0

-MAXLEN-2000000000 -USPR-US09748451.ecgn.1.1.57.runat.21052003.153830.16347

-NCPI-6 -ICPI-3 -NO.MMAP -LARGEDUERY -NEG.SCORES-0 -WAIT -DSPELOCK-100

-IONLOG -DEV.TIMEOUT-120 -MARR.TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5

-FGAPOP-6 -FGAPEXT-7 -FGAPOP-10 -FGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database:

Published.Applications.NA.*

1: /cgna2.6/ptodata/1/pubpna/US07_PUBCOMB.seq.*

2: /cgna2.6/ptodata/1/pubpna/US06_PUB_PUB.seq.*

3: /cgna2.6/ptodata/1/pubpna/US06_PUBCOMB.seq.*

4: /cgna2.6/ptodata/1/pubpna/US07_NEW_PUB.seq.*

5: /cgna2.6/ptodata/1/pubpna/US08_NEW_PUB.seq.*

6: /cgna2.6/ptodata/1/pubpna/US08_NEW_PUBCOMB.seq.*

7: /cgna2.6/ptodata/1/pubpna/US09_NEW_PUB.seq.*

8: /cgna2.6/ptodata/1/pubpna/US09_NEW_PUBCOMB.seq.*

9: /cgna2.6/ptodata/1/pubpna/US10_NEW_PUB.seq.*

10: /cgna2.6/ptodata/1/pubpna/US10_NEW_PUBCOMB.seq.*

11: /cgna2.6/ptodata/1/pubpna/US10_NEW_PUBCOMB.seq.*

12: /cgna2.6/ptodata/1/pubpna/US10_NEW_PUBCOMB.seq.*

13: /cgna2.6/ptodata/1/pubpna/US10_NEW_PUBCOMB.seq.*

14: /cgna2.6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result	Score	Query	Match	Length	ID	Description
1	1789	100.0	2839	10	US-09-748-451-1	Sequence 1, Appl1
2	102.5	3.7	1766	9	US-10-242-963-11	Sequence 11, Appl1
3	102.5	3.7	1862	9	US-10-161-803-29	Sequence 29, Appl1
4	102.5	5.7	1864	9	US-09-974-298-43	Sequence 43, Appl1

5	102.5	5.7	1892	12	US-10-044-090-793	Sequence 793, App
6	102.5	5.7	3405	9	US-09-974-298-44	Sequence 44, Appl1
7	102.5	5.7	3405	9	US-09-981-353-36	Sequence 36, Appl1
8	99.5	5.6	1218	9	US-10-043-487-101	Sequence 101, App
9	99.5	5.6	2209	9	US-10-161-803-30	Sequence 30, Appl1
10	99.5	5.4	69327	10	US-09-777-921A-3	Sequence 3, Appl1
11	97.5	5.4	1749	9	US-09-923-779-149	Sequence 149, App
12	96.5	5.4	2673	9	US-10-060-425-7	Sequence 7, Appl1
13	95.5	5.3	1922	9	US-09-995-898A-32	Sequence 32, Appl1
14	95.5	5.2	1225	10	US-09-216-393-284	Sequence 284, App
15	92.5	5.2	1225	10	US-09-216-393-284	Sequence 3985, App
16	90	5.0	1282	9	US-09-764-881-988A	Sequence 988A, App
17	90	5.0	1282	9	US-09-764-881-988A	Sequence 988, App
18	90	5.0	1467	10	US-09-834-975-958	Sequence 958, App
19	90	5.0	11358	9	US-10-091-572-887	Sequence 887, App
20	90	5.0	11358	9	US-09-764-881-9340	Sequence 9340, App
21	90	5.0	11600	9	US-10-091-572-888	Sequence 888, App
22	90	5.0	11600	9	US-09-764-881-9341	Sequence 9341, App
23	89.5	5.0	11770	9	US-10-097-065-121	Sequence 121, App
24	89	5.0	1689	10	US-09-748-451-3	Sequence 3, Appl1
25	89	5.0	3309400	9	US-09-738-626-1	Sequence 1, Appl1
26	88.5	4.9	735	10	US-09-770-149-82	Sequence 82, Appl1
27	88.5	4.9	1074	10	US-09-974-300-2794	Sequence 2794, App
28	88.5	4.9	2084	10	US-09-822-849A-526	Sequence 526, App
29	88	4.9	4235	10	US-09-764-877-3368	Sequence 3368, App
30	88	4.9	10955	10	US-09-070-927A-236	Sequence 236, App
31	87.5	4.9	579	10	US-09-974-300-2935	Sequence 2935, App
32	87.5	4.9	606	10	US-10-165-800-10	Sequence 10, Appl1
33	87.5	4.9	1284	9	US-09-974-300-3376	Sequence 3376, App
34	87	4.9	1728	10	US-09-925-300-555	Sequence 555, App
35	87	4.9	2602	9	US-09-974-298-45	Sequence 45, Appl1
36	87	4.9	17239	9	US-10-091-504-2326	Sequence 2326, App
37	87	4.9	17239	10	US-09-764-869-2326	Sequence 2326, App
38	86.5	4.8	657	9	US-10-165-800-12	Sequence 12, Appl1
39	86	4.8	5387	9	US-10-001-873-22	Sequence 22, Appl1
40	86	4.8	868	9	US-09-938-842A-1748	Sequence 1748, App
41	85.5	4.8	2019	9	US-10-097-340A-17	Sequence 17, Appl1
42	85.5	4.8	2274	9	US-10-097-340-17	Sequence 17, Appl1
43	85.5	4.8	14886	9	US-10-203-224-1	Sequence 3209, App
44	85	4.8	353	9	US-09-918-995-32209	Sequence 26, Appl1
45	85	4.8	1458	9	US-09-808-602-26	

ALIGNMENTS

RESULT 1
US-09-748-451-1
Sequence 1, Application US/09748451
Patent No. US20010011078A1
GENERAL INFORMATION:
APPLICANT: WANG, XIAODONG
APPLICANT: LIU, XUESONG
TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN APOPTOSIS
FILE REFERENCE: US05:546USD1
CURRENT APPLICATION NUMBER: US/09/748,451
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 09/061,702
PRIOR FILING DATE: 1998-04-16
NUMBER OF SEQ ID NOS: 21
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO. 1
LENGTH: 2839
TYPE: DNA
ORGANISM: Homo sapiens
US-09-748-451-1

Alignment Scores:
Pred. No.: 5.83e-223
Score: 1789.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 10
Gaps: 0

US-09-748-451-2 (1-338) x US-09-748-451-1 (1-2839)

QY 1 MetLeuGlnYsPProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
 DB 132 ATGCTCCAGAAAGCCCAAGAGGTGACAGCTGGGGCCCTGCCAGCCGAGAAATTGGCC 191
 QY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40
 DB 192 GTGGCTGGCCGAGCTGCGCAGAGAGCTGCGCAAGGGCTGTCCGCTTCACAGTCCCT 251
 QY 41 GlnArgGlySerArgLysLeuGlyLeuTyrGlnAspGlyThrGluLeuThrGluAspTyrPhe 60
 DB 252 GAGCGCGGTTCGCGCTGCTGCTACGAGAGATGCGAGGAGCGAGGAAATTAATCTC 311
 QY 61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaTyrGlnGly 80
 DB 312 CCCAGCTGTCCCGAACAACGCCAGACTGCTGCTCACTCGGGCCAGCGCTGGCAGGGC 371
 QY 81 TyrValSerAspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
 DB 372 TATGTGAGCGACATCAGAGCGCTCTCTACATTCATTCACAGCCACAGTGGGCTCATC 431
 QY 101 GlnAlaIaGlnGlnLeuLeuCysAspGlnGlnAlaProGlnArgGlnArgLeuAla 120
 DB 432 CAGGCGCCGACAGAGCTGCTGTGTGATGAGCAGGCCCCACAGAGGAGGCTGCTGCT 491
 QY 121 AspleuLeuHisAsnValSerGlnAsnIleAlaIaGluThrArgAlaGluAspProPro 140
 DB 492 GACCTCTGCAACAGTCAGGCGACAGACATGCGGGCCGAGACCGGGCTGAGAGACCGCGC 551
 QY 141 TyrPheGluGlyLeuGlnSerArgPheGlnSerLysSerGlyTyrLeuArgTyrSerCys 160
 DB 552 TGGTTTAAAGGCTTGAAGTCCCGATTCAGAGCAAGCTGCTATCTGAGATACAGCTGT 611
 QY 161 GluSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
 DB 612 GAGAGCGCGATCCGAGATTCCTGAGAGAGGTGAGCTCTTACCCCTCCACAGTGGGTGCG 671
 QY 181 GlnAlaGlnGlnGlnLeuLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
 DB 672 GAGGCTCAGAGAGAAATTCCTGCGGGTCTCGGCTCCATGTCAGAGGCTCGGTCATG 731
 QY 201 GlnGlyArgGlnSerTyrPheAspArgGlyAlaLysGlySerArgLeuGlyCysThrPro 220
 DB 732 CAGTACATGCGACACTTCTGACAGAGAGCCAAAGGGCGGACCGCTCTGACACCG 791
 QY 221 GlnGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 DB 792 GAAAGCGGTCTCTCTGCGAGGCTCTTTCACATGGACAGCTCTTATCAAGACACTCC 851
 QY 241 IleAsnProTyrSerAsnArgGlnSerArgIleLeuPheSerThrTyrPasnLeuAspHis 260
 DB 852 ATCAACCCCTACAGTACAGAGAGAGAGCATCTCTTCAGACACCTGAAACCTGATCAC 911
 QY 261 IleIleGluLysLysAspGlnIleIleProThrLeuValGluAlaIleLysGlnGlnAsp 280
 DB 912 ATATATACAAAGAAAGCCACCATCATCTTCTACACTGTGGAAAGCAATTAAGAAACAAGAT 971
 QY 281 GlnArgGlnValAspTyrPheTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeuLys 300
 DB 972 GGAAGAAAGAGTGGAGTATTTTATGGCTGCTTTTACTCTAGAGAACTTAA 1031
 QY 301 LeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSerArg 320
 DB 1032 CTAGTGCACATGCTGCTCAATAGAAACACCAACAGCTCACTGAGACCCAGAGCA 1091
 QY 321 IleGlyTyrProGlnThrArgLeuLysArgLysGlnProValArgLysArgGln 338
 DB 1092 ATCTACAAACCCAGACAGAGGTGAAGCGGAAGAGCGCTGTGCGAAACCGCCAG 1145
 RESULT 2
 US-10-242-943-11

; Sequence 11, Application US/10242943
 ; Publication No. US20030087412A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Nandabalan, Krishnan
 ; APPLICANT: Schulz, Vincent P.
 ; APPLICANT: Yang, Melja
 ; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
 ; FILE REFERENCE: 15966-521 NIK1 protein complexes
 ; CURRENT APPLICATION NUMBER: US/10/242,943
 ; PRIORITY FILING DATE: 2002-09-13
 ; PRIOR APPLICATION NUMBER: US/09/167,206
 ; PRIOR FILING DATE: 1998-10-06
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 11
 ; LENGTH: 1766
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (44)..(1441)
 US-10-242-943-11

Alignment Scores:
 Pred. No.: 0.00537 Length: 1766
 Score: 102.50 Matches: 77
 Percent Similarity: 34.42% Conservative: 50
 Best Local Similarity: 20.87% Mismatches: 128
 Query Match: 5.73% Indels: 115
 Gaps: 16

US-09-748-451-2 (1-338) x US-10-242-943-11 (1-1766)

QY 4 LysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyValAlaGly 23
 DB 546 AAGCCCGCGTGG---AGGTGAGCGCGCAGCAACCTGGCCGAGAGATCATGCGCTCGGG 602
 QY 24 ArgSerCysGlnGlnValLeuArgLysGlyCysLeuArgPheGlnLeuProGluArgGly 43
 DB 603 AGAAATTCAGAG---
 QY 44 SerArgLeuGlyLeuTyrGlnAspGlyThrGluLeuThrGluAspTyrPheProSerVal 63
 DB 615 ---AGGAGATGCTTCAGAGAGAG---GAAAGCGAAACACCTGCAATCTTTCAGACAGAT 670
 QY 64 ProAspAsnAlaGlnLeuValLeuLeuThrLeuGlnGlnAlaTyrGlnGlyTyrValSer 83
 DB 671 GTTACATAGTGGTCTGCGACGCTTGACCTTGAAAGCAAGAGAAATCTTTCAGAGAA 730
 QY 84 AspleuArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAlaAla 103
 DB 731 GAGATT---GCCTTTGGAGAAACCTCCACAGAGAAACCAAGAGAGAGCTGACAGCTCAG 787
 QY 104 ---GlnGlnLeuLeuCysAspGlnGlnAlaProGlnArgGlnArgLeu 118
 DB 788 ATTACAGAACAGCATGCTCAATATCATGATGTGATGTTTCCAAAGCTGACCTCAGCGCTGCC 847
 QY 119 LeuAlaAspleuLeuHisAsnValSerGlnAsnIleAlaIaGluThrArgAlaGluAsp 138
 DB 848 CTGGCTGACGTAGTCAGCAATAT---GAAAGTGTGGCTGCGCAAGACCTGCAAGAGCA 904
 QY 139 ProThrTyrPhe-----Glu 143
 DB 905 GAAAGAAAGTGTCAATCCAAATCCAAATGCTGACCTCTGAGCTGCCAACCGGAAACAATGAC 964
 QY 144 GlyLeuGlnSerArgPheGlnSerLysSerGlyTyrLeuArgTyr----- 158
 DB 965 GCCCTGGCGCCAGCAAGCAGAGAGTCCACTGATGACGAGACAGGTCAGCTCCCTCACC 1024
 QY 159 -----SerCysGlnSerArgIleArg----- 165
 DB 1025 TGTGAGTGGATGCGCTTAAAGAGAACCAATGATGCTCGAAGAGCCAGATCGTGAATG 1084


```

: RESULT 4
: US-09-974-298-43
: Sequence 43, Application US/09974298
: Patent No. US20020156263A1
: GENERAL INFORMATION:
: APPLICANT: Chen, Huel-Mei
: TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
: FILE REFERENCE: PA-0037 P
: CURRENT APPLICATION NUMBER: US/09/974,298
: CURRENT FILING DATE: 2001-10-04
: PRIOR APPLICATION NUMBER: 60/238,331
: PRIOR FILING DATE: 2000-05-10
: NUMBER OF SEQ ID NOS: 194
: SOFTWARE: PERL Program
: SEQ ID NO 43
: LENGTH: 1864
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc.feature
: OTHER INFORMATION: Inocyte ID NO. US20020156263A1 127987.19
: 9-974-298-43

Alignment Scores:
Pred. No.: 0.00584 Length: 1864
Score: 102.50 Matches: 77
Percent Similarity: 34.42% Conservative: 50
Best Local Similarity: 20.87% Mismatches: 128
Query Match: 5.73% Indels: 115
DB: Gaps: 16

US-09-748-451-2 (1-338) x US-09-974-298-43 (1-1864)

OY 4 LysProlySservAllysLeuArgAlaLeuArgSerProAlglySPheGlyValAlaGly 23
Db 642 AAGCCCGGCTG--AGGTGGAGGCGCCAACTGGCCGAGACATCATCGCCCTCCGG 698
24 ArgSerCySGInglValLeuArgLysGlyCysLeuArgPheGlnLeuProGluArgLys 43
Db 699 AGAAATTCAGG----- 710
OY 44 SerArgLeuCySLeuTyrgLAspGlyThrgLLeuThrgLAspTyrgPheProSerVal 63
Db 711 --AGGAGATGCTTCACAGAGAG--GAAGCCGAAACACCCTGCATCTTTCAGACAGAT 766
OY 64 ProAspAsnAlaGlnLeuValLeuLeuThrgLLeuGlnAlaIrpGlnGlyTyrgValSer 83
Db 767 GTTGACATGAGCTGCTGGACGCTGTGCATTCGAAACCAAGTGAATCTTTCGACAGA 826
OY 84 AspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAlaIa 103
Db 827 GAGATT---GCCTTTTCGAAACCTCCACAGAGAAAGAAATCCACAGAGCTCAGGCTCAG 883
104 -----GlnGlnLeuLeuCySAspGlnAlaProGlnArgGlnArgLeu 118
Db 884 ATTCAGCAACGATGCTCCAAATGATGATGTTCACAGCCTGACCTCAGCGCTGCC 943
OY 119 LeuAlaAspLeuHisAsnValSerGlnAsnIleAlaGlnThrArgAlaGlnAsp 138
Db 944 CTGGCTGACGTACGTACCAATAT--GAAAGTGTGCTGCCAAGACCTCAGAGAGCA 1000
OY 139 ProProTrpPhe-----Glu 143
Db 1001 GAAGAATGTCACAAATCCAAATTGTCGTACCTCTGAGGCTGCCAACCGAACAATGAC 1060
OY 144 GlyLeuLusArgPheGlnSerLysSerGlyTyrgLeuArgTyrg----- 158
Db 1061 GCCCTGGCCAGCAAGACAGAGTCCACTGATACCGAGACAGAGTCCATCCCTCAC 1120
OY 159 -----SerCySgluSerArgIleArg----- 165
Db 1121 TGTCAGATGATGCCCTTAAAGAACCAATGATGTCCTGTCGACCAACCAATGCTGGAATG 1180

```

QY	166	-----SetrTyrLeuArgIuValSerSetrTyrProserThrValGlyAlaIuValGln	183
QY	1181	GAAGACAACTTGGCCCTTGAAGCTGCTCACTACCCACAGACTAATTGGC---CGCCTGCAG	1237
QY	184	GIuGIuPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsn	203
Db	1238	GATGAGATTCAGAAATATGAAGAAGAAATGGCTGCTGACCTTCGTGAATACCA-----	1291
QY	204	GlySerTyrPheAspArgGlyAlaValGlySerArgLeuGlyThrProGluGlyTyr	223
Db	1292	-----GACCTG	1297
QY	224	PheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSerIleAsnPro	243
Db	1298	CTCAAACTTAAAGTATGGCCCTTGAACATTTGAAATGGACCTACAGAAAGCTGCTG-----	1351
QY	244	TyrSerAsnArgGlySerArgIleLeu-----PheSerThr-----	255
Db	1352	---GAAGGGGAGAGAGCAGGATTTCTGCTCCCTTCCAAACTTTCTCCTCCAGACTG	1408
QY	256	-----TTPAsnLeuAspHisIle-----IleGluTyrSerArgThrIleIle	269
Db	1409	AGGGAACAACTAATCTGGATTCACCTCCCTGCTGTGATACCCACATCAAAAGAGCACTT---	1465
QY	270	ProThrLeuValGluAlaIleLysGluGlnAspGlyArgIuValAsp-----	285
Db	1466	-----CTGATTAAAGACGGTTGGAACACTAGAGATGGACAGGTTATCAACAGAACTTCTCAG	1519
QY	286	-----TTPGluTyrPheTyrGlyLeuLeuPhe-ThrSerGln	297
Db	1520	CATCAAGATGACCTTGATTAATAATTGCACACACTAGTGCAGCAATATATTACAGCAA	1579
QY	297	uAsnLeuLysLeuValHisIleVal	305
Db	1580	GATATAAAAGAAATCATCATATCTTA	1604
RESULT 5	US-10-044-090-793		
	Sequence 793, Application US/10044090		
	Patent No. US20020137081A1		
	GENERAL INFORMATION:		
	APPLICANT: Olga Bandman		
	TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION		
	FILE REFERENCE: PA-0028 US		
	CURRENT APPLICATION NUMBER: US/10/044, 090		
	CURRENT FILING DATE: 2002-01-09		
	NUMBER OF SEQ ID NOS: 850		
	SOFTWARE: PERL Program		
	SEQ ID NO 793		
	LENGTH: 1892		
	TYPE: DNA		
	ORGANISM: Homo sapiens		
	FEATURE:		
	NAME/KEY: misc_feature		
	OTHER INFORMATION: Incyte ID No. US20020137081A1 2812002CBI		
	US-10-044-090-793		
Alignment Scores:			
Pred. No.:	0.00598	Length:	1892
Score:	102.50	Matches:	77
Percent Similarity:	34.42%	Conservative:	50
Best Local Similarity:	20.87%	Mismatches:	128
Query Match:	5.73%	Indels:	115
DB:	12	Gaps:	16
US-09-748-451-2 (1-338) x US-10-044-090-793 (1-1892)			
QY	4	IysProlsSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyValAlaGly	23
Db	657	AAGCCCGGCTGC---AGTGGAGAGCGGACAACTGGCCGAGACATCATGCGCTCCGGG	713
QY	24	ArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuProGluArgGly	43

```

Db 714 AGAATTGTCAGG-----725
OY 44 SerArgLeuCyLeuTYrGluAspGlyThrGluLeuThrGluAspTYrPheProSerVal 63
Db 726 ---AGAGATGCTTCAGACAGAG--GAAGCCGAAACACCTGCATCTTTCAGACAGAT 781
OY 64 ProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaIleArgGlnGlyValSer 83
Db 782 GTTGACCAATGGCTCTCGCACGCTTGACCTTGAAACGGAAGTGGATCTTTCAGAA 841
OY 84 AspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAla 103
Db 842 GAGATT---GCCTTTTAAAGAACTCCAGAAAGAAATCCAGAGCTCAGGCTCAG 898
OY 104 -----GlnGlnLeuLeuCyAspGlnGlnAlaProGlnArgGlnArgLeu 118
Db 899 ATTTCAGGAACAGCATGTCCTCAATGATGTGATGTTTCCAAAGCTGACCTCAGCTGCC 958
OY 119 LeuAlaAspLeuLeuHisAsnValSerGlnAsnIleAlaIleGluThrArgAlaGluAsp 138
Db 959 CTGCTGACGTACGTACAGCAATAT--GAAGTGTGCTGCCAAGACCTGCAGAGAGCA 1015
OY 139 ProProTrpPhe-----Glu 143
Db 1016 GAAGAATGTGTACAAATCCAAATTGCTGACCTCTGAGGCTGCCAAGCAGACAAATGAC 1075
OY 144 GlyLeuGluSerArgPheGlnSerIleSerGlyTYrLeuArgTYr-----158
Db 1076 GCCCTGCCAGCAAGCAAGCAAGTCACAGTACCCGAGCAGAGTCCATCCCTCAC 1135
OY 159 -----SerCyGluSerArgIleArg-----165
Db 1136 TGTGAAGTATGATCCCTTAAGAACCAATGATGATCCCTGAAAGCAGAGTGGTGAATG 1195
OY 166 -----SerTYrLeuArgGluValSerSerTYrProSerThrValGlyAlaGluAlaGln 183
Db 1196 GAAGAACTTTCGCCGTGAAGCTGTACTACCAAGACACTAATGGC---CCCTGCAG 1252
OY 184 GlnGluPheLeuArgValLeuGlySerMetCyGlnArgLeuArgSerMetGlnTYrAsn 203
Db 1253 GATGAGATTCAGAAATATGAAGAGAAATGGCTGCTGACCTTCGTGATACCA-----1306
OY 204 GlySerTYrPheAspArgGlyAlaIleGlyGlySerArgLeuCyThrProGlnGlyTYr 223
Db 1307 -----GACCTG 1312
OY 224 PheSerCyGlnGlyProPheAspMetAspSerCyLeuSerArgHisSerIleAsnPro 243
Db 1313 CTCATATGTTAAAGATGCGCTTCGACATTGAGATTGCCACTACAGGAAGCTGCTG--1366
OY 244 TYrSerAsnArgGluSerArgIleLeu-----PheSerThr-----255
Db 1367 ---GAAGCGAGAGAGAGAGATTTCTCTGCTCTTCCTCAAACTTTTCCCTCGAAGCTG 1423
OY 256 -----TrpAsnLeuAspHisIle-----IleGlnIleGlySerArgThrIleIle 269
Db 1424 AGGAAACATATCTGATTCACCTCTGCTGTGATACCCACTCAAAAAGGACACTT---1480
OY 270 ProThrLeuValGlnAlaIleGlyGlnIleAspGlyArgGluValAsp-----285
Db 1481 -----CTGATTAAAGCGTTGAAGACTAGAGATGACAGCTATATACAGAAACTTCTCAG 1534
OY 286 -----TrpGluTYrPheTYrGlyLeuLeuPhe-ThrSerGln 297
Db 1535 CATCAGATGACCTTGATTAATAATTCACACACTCAGTGCAGCAATATATTACAGAGAA 1594
OY 297 uAsnLeuValSerValHisIleVal 305
Db 1595 GATTAAAAAGAAATCATATCTTA 1619
OY 166 -----SerTYrLeuArgGluValSerSerTYrProSerThrValGlyAlaGluAlaGln 183

```

RESULT 6
US-09-974-298-44

```

/ Sequence 44, Application US/09974298
/ Patent No. US20020156263A1
/ GENERAL INFORMATION:
/ APPLICANT: Chen, Huel-Mei
/ TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
/ FILE REFERENCE: PA-0037 P
/ CURRENT APPLICATION NUMBER: US/09/974,298
/ PRIOR FILING DATE: 2001-10-04
/ PRIOR APPLICATION NUMBER: 60/238,331
/ NUMBER OF SEQ ID NOS: 194
/ SOFTWARE: PERL Program
/ SEQ ID NO 44
/ LENGTH: 3405
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No. US20020156263A1 1397781.7
/ US-09-974-298-44

Alignment Scores:
Pred. No.: 0.015 Length: 3405
Score: 102.50 Matches: 77
Percent Similarity: 34.42% Conservative: 50
Best Local Similarity: 20.87% Mismatches: 128
Query Match: 5.73% Indels: 115
DB: 9 Gaps: 16

US-09-748-451-2 (1-338) x US-09-974-298-44 (1-3405)
OY 4 LysProIleSerValIleLeuArgAlaLeuArgSerProArgIlePheGlyValAlaGly 23
Db 784 AACCCGCCGTCG---AGGTGAGCGCGCAACCACTGGCGGAGGACATCATCGCCCTCGGG 840
OY 24 ArgSerCyGlnGlnValLeuArgGlySerIleLeuArgPheGlnLeuProGlnArgGly 43
Db 841 AGAATTCGACG-----852
OY 44 SerArgLeuCyLeuTYrGluAspGlyThrGluLeuThrGluAspTYrPheProSerVal 63
Db 853 ---AGAGATGCTTCAGACAGAG--GAAGCCGAAACACCTGCATCTTTCAGACAGAT 908
OY 64 ProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaIleArgGlnGlyValSer 83
Db 909 GTTGACCAATGGCTCTCGCACGCTTGTGACCTTGAAACGGAAGTGGATCTTTCAGAA 968
OY 84 AspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAla 103
Db 969 GAGATT---GCCTTTTGAAGAACTCCAGAAAGAAATTCAGAGAGCTCAGGCTCAG 1025
OY 104 -----GlnGlnLeuLeuCyAspGlnGlnAlaProGlnArgGlnArgLeu 118
Db 1026 ATTTCAGGAACAGCATGTCCTCAATGATGTGATGTTTCCAAAGCTGACCTCAGGCTGCC 1085
OY 119 LeuAlaAspLeuLeuHisAsnValSerGlnAsnIleAlaIleGluThrArgAlaGluAsp 138
Db 1086 CTGCTGACGTACGTACAGCAATAT--GAAGTGTGCTGCCAAGAACTGCAGAGAGCA 1142
OY 139 ProProTrpPhe-----Glu 143
Db 1143 GAAGAATGTGTACAAATCCAAATTGCTGACCTCTGAGGCTGCCAAGCAGGAACAATGAC 1202
OY 144 GlyLeuGluSerArgPheGlnSerIleSerGlyTYrLeuArgTYr-----158
Db 1203 GCCCTGCCAGCAAGCAGAGATGCCACTGATAGTCAGGACAGTGCAGTCCCTCAGC 1262
OY 159 -----SerCyGlnSerArgIleArg-----165
Db 1263 TGTGAAGTATGATCCCTTAAGCAACCAATGATGCTGGAAGCGCAGATGCTGGAATG 1322
OY 166 -----SerTYrLeuArgGluValSerSerTYrProSerThrValGlyAlaGluAlaGln 183

```

```

Db      1333 GAAGAGAACTTGCCTTGAAGCTCTCAATACCAAGACACTATTGGC---CGCCTGAG 1379
Oy      184  GUGUGUPELEUAAGVALLEUGLISERKETCYSGINALEUGAUGSERMELGINTYRASN 203
Db      1380 GATGAGATTCAAGATATATAGAGAGAAATAGCTCTCACCCTTCGGAATACCA----- 1433
Oy      204  GLYSERTYRPHASPARGLYALALYSGLYSERATGLEUCYSTHPRGLOUGLYTRP 223
Db      1434 -----GACCTG 1433
Oy      224  PHESECYSGINGLYPROPHASPMETASPSERTCYLSEUSERARGHISERILEANPRO 243
Db      1440 CTCAATCTTAAGATWGCCCCCTTGACATTTGAGATTCGACACCTACAGSAACTGCTG----- 1493
Oy      244  TYRSEIRANALRGJUSERATRIILEU-----PHESETRH----- 255
Db      1494 --GNAGCGGAGAGAGCAGATTTCTCTGCCCTTCACAACTTTCTCCCTCGAACCCTG 1556
Oy      256 -----TRPANELUASPHISILE-----ILEGILYSLYARGTHRIELLE 269
Db      1551 AGGGAACCTAATCTGGAATCACTCCCTCTGTTGATACCACCTCAAAAAGACACTT--- 1607
Oy      270  PROTHLEUVALGLUALAILELYSGIUGINASPGIYARGJUALASP----- 285
Db      1608 -----CTGATTTAGACCGCTTGAAACTAGAGATGACAGGTTATCAACGAACCTTCAG 1666
Oy      286 -----TRPGLUTYRPHERYGLYLEUPEHPE-THSERGI 297
Db      1662 CATCAGATGACCTTGAAATAAATATGCACACACTCACTGACGACAAATATATACAGCAA 1721
Oy      297  UASNLEUVALYSLEUVALHISILEVAL 305
Db      1722 GAATRAAAAAGAAATTCATATCTTA 1746

RESULT 7
US-09-981-353-36
Sequence 36, Application US/09981353
Patent No. US20020160382A1
GENERAL INFORMATION:
APPLICANT: Lasek, Amy W.
TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
FILE REFERENCE: PA-0038 US
CURRENT APPLICATION NUMBER: US/09/981.353
CURRENT FILING DATE: 2001-10-11
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 36
LENGTH: 3405
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020160382A1 1397781.7
-981-353-36

Alignment Scores:
Pred. No.: 0.015 Length: 3405
Score: 102.50 Matches: 77
Percent Similarity: 34.42% Conservative: 50
Best Local Similarity: 20.87% Mismatches: 128
Query Match: 5.73% Indels: 115
DB: Gaps: 16
US-09-748-451-2 (1-338) x US-09-981-353-36 (1-3405)
Oy      4  LYSPPOLYSSEVALYSLEUAARGALALEUAARGSERPROATGGLYSPHGLYVALALGLY 23
Db      784  AAGCCCGGCTG---AGGTGGAGCGCGCAACCTGGCGGAGAGACATCATGCGCTCCGGG 840
Oy      24  ARGSEYSGINGIUALLEUAARGLSLYSYSEUAARGPHEGINLEUPROGLUARGLY 43
Db      841  AGAAATGCGAG----- 852

```

Oy	44	SerTyrLeuGlySerLeuTyrGlyAlaSpolYThrGlnGluLeuThrGlnAlaSerTyrPheProSerVal	63
Db	853	---AGGAGATCCTTCAGAGAGAG- GAAGCCAAAACCCCTGCATCTTTAGACAGGAT	908
Oy	64	ProAspAsnAlaGluLeuValLeuLeuThrLeuGlnAlaIleArgGlnGlyTyrValSer	83
Db	909	GTTCAGCATCGGCTCGTCGACGCGTTCGACCTTGAAACCAAGTGGAACTTTTGCAACGA	968
Oy	84	AspIleArgValPheLeuSerValPheHisGluProGlnValGlyLeuIleGlnAlaIa	103
Db	969	GAGATT---GCCTTTGGAAAGAACTCCACAGAGAAATCCAGAGACTCAGAGCTCAG	1025
Oy	104	-----GlnGlnLeuLeuGlySerGluGlnAlaIleProGlnIleArgGlnIleArgLeu	118
Db	1026	ATTGAGGAACAGCATGTCCAAATGATGTGATGTTCCTCAAGCCTGACCTCAGGCTCC	1085
Oy	119	LeuAlaSpLeuLeuHisAsnValSerGlnAsnIleAlaIleAlaGluThrArgAlaGluAsp	138
Db	1086	CTGGGTGACGTCAGCGACAGCAATAT---GAAGTGTGGCTGGCAAGAACTCAGAGAGCA	1142
Oy	139	ProProTyrPhe-----	143
Db	1143	GAAGAAATGCTCAAAATCCAAAGTTTGTCTGACCTCTGTGAGGCTGCCAACCGGAACAAATGAC	1202
Oy	144	GlyLeuGlnSerArgPheGlnSerTyrSerGlyTyrLeuArgTyr-----	158
Db	1203	GCCCTGCGCCACGGCAAGACAGGAGTCCACTGATGACGGAACAGAGTCCACTCCCTCACCC	1265
Oy	159	-----SerGlyGlnSerArgIleArg-----	165
Db	1263	TGTGAAGTGATGCCCTTAAAGAAACCAATGATGATCCCTGGAACGCCAGATCGTGAATG	1322
Oy	166	-----SerTyrLeuAlaGlyValSerSerTyrTyrProSerThrValGlyAlaGlnAlaGln	183
Db	1323	GAAGAGAACTTTGCCGTGAAGCTGCTACATCAACAGACACTATTTGCG---CGCCTGAG	1379
Oy	184	GluGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsn	203
Db	1380	GATGAGATTCAGATATATGAGAGAGAAATGAGCTGCTGACCTTCGTGATACCA-----	1433
Oy	204	GlySerTyrPheAspArgGlyAlaValGlyGlySerArgLeuGlySerThrProGluGlyTyr	223
Db	1434	-----GACCTG	1439
Oy	224	PheSerCysGlnGlyProPheAspMetAspSerGlySerValSerArgHisSerIleAsnPro	243
Db	1440	CTCATATGTTAAGATGGCCCTTGACATTTGAGATGGCCACACTACAGGAAGCTGCTG-----	1493
Oy	244	TyrSerAsnArgGlnSerArgIleLeu-----PheSerThr-----	255
Db	1494	---GAAGCGGAGAGAGAGAGATTTCTCTGCTCTCCAAACTTTCTCTCCCTGAACTG	1550
Oy	256	-----TrpAsnLeuAspHisIle-----IleGluValGlySerArgThrIleIle	269
Db	1551	AGGGAACCTATTCGATTCACCTCCCTGTTATACCCACTGAAAAAGACACTT---	1607
Oy	270	ProThrIleValGluAlaIleLeuGlyGlnAspGluAspArgGluValAsp-----	285
Db	1608	-----CTGATTAAAGACGGTTGAAACTGAGATGACAGGTTATCAACGAAACTTTCAG	1661
Oy	286	-----TrpIuTyrPheTyrGlyLeuLeuPhe---ThrSerGln-----	297
Db	1662	CATCAACATGACCTTGAAATAAAAATTCGACACACTCACTGACGAGCAATATATTACAGCAA	1721
Oy	297	uAsnLeuValHisIleVal	305
Db	1722	GAATTAATAAAGAAATCATATCTTA	1746

OY 84 asplleatgargphleuseSerlaPhehisGlnProInValGlyLeuIleGlnIlela 103
 Db 770 GAGATT---GCCTTTGGAAAGAACTCCACGAGGAAATCCAGGAGCTGCAG 826
 OY 104 -----GInGlnLeuLeuCysAspGlnInalProGlnatrgInatrgLeu 118
 Db 827 ATTGAGGAACGACATGTCCAATTCGATGATGATGTTTCCAAAGCTGACACGAGGCTGCC 886
 OY 119 LeuAlaSplLeuLeuHisAsnValSerGlnAsnIleAlaGluThrArgAlaGluAsp 138
 Db 887 CTGGGTACGTCAGTCACGACAAATAT- GAAGAAGTGGCTGCCAAGAACCTGCAGAGAGCA 943
 OY 139 ProProtrpPhe-----Glu 143
 Db 944 GAAAGATATGTCACAAATCCAAAGTTTGTCTGACCTCTCTAGGCTGCCAACCGAACAATGAC 1003
 OY 144 GlyLeuGlnSerArgPheGlnSerLysSerGlyTyrLeuArgTyr----- 158
 Db 1004 GCCCTGGCGCCAGCAAGCAGAGAGTCCACTGAGACCGAGAGGACAGTCCCTCAC 1063
 OY 159 -----SerCysGlnSerArgIleArg----- 165
 Db 1064 TGTGAGTGTGATGCCCTTAAAGCAACATGAGTCCCTGGAGAGCCAGATGCGTGAATG 1123
 OY 166 -----SerTyrLeuArgGluValSerSerTyrProSerThrValGlyAlaGluAlaGln 183
 Db 1124 GAAGAGAAACTTGGCGCTTGAAGAGCTGCTACTACCAAGACACTTTGGC---CCCTTCAG 1180
 OY 184 GlnGlnPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsn 203
 Db 1181 GATAGAGTTGACGAATGAAGAGGAAATGCTGCTCCACTGCTGATACCA----- 1234
 OY 204 GlySerTyrPheAspArgGlyAlaLeuGlyGlySerArgLeuSerThrProGlnGlyTyr 223
 Db 1235 -----GACTTG 1240
 OY 224 PheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSerIleAsnPro 243
 Db 1241 CTCAAATGTTAAGATGAGCCCTTGACATGTAGATTTGCCACCTACAGAGAACTGCTG----- 1294
 OY 244 TyrSerAsnArgGlnSerArgIleLeu-----PheSerThrTrpAsnLeu 258
 Db 1295 ---GAAGCGCAGACAGACAGATTTCTCTGCTCTTCCAAACTTTTCCCTCCGAACCTG 1351
 OY 259 -----AspHisIleIleGlu-----LysLysArgThrIleLeuProThrLeuValGlu 274
 Db 1352 AGGGGTAAAGCATTTATTTCCTTTAGAGAAAACGACAGCTGCTTTAACCACTGCTTT 1411
 OY 275 AlaIleLysGlnGlnAspGlyArgGluValAspTrpGluTyrPheTyrGlyLeuLeuPhe 294
 Db 1412 ATGTCAAGCAT-----TCATTTTATTAGATAT----- 1441
 OY 295 ThrSerLysnLeuLysLeuValHisIle 304
 Db 1442 -----CTGAATAAAATGCCATATA 1459
 RESULT 10
 US-09-777-921A-3
 : Sequence 3, Application US/09777921A
 : Patent No. US20020115136A1
 : GENERAL INFORMATION:
 : APPLICANT: MERKULOV et al.
 : TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
 : TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
 : TITLE OF INVENTION: AND USUS THEREOF
 : FILE REFERENCE: CL001103
 : CURRENT APPLICATION NUMBER: US/09/777,921A
 : NUMBER OF SEQ ID NOS: 126
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 3
 : LENGTH: 69327

:	TYPE:	DNA
:	ORGANISM:	Homo sapiens
:	FEATURE:	
:	NAME/KEY:	misc_feature
:	LOCATION:	(1)...(69327)
:	OTHER INFORMATION:	n = A,T,C or G
:	US-09-777-921A-3	
 Alignment Scores:		
Pred. No.:	4.08	Length: 69327
Score:	99.50	Matches: 86
Percent Similarity:	39.69%	Conservative: 66
Best Local Similarity:	22.45%	Mismatches: 123
Query Match:	5.56%	Indels: 108
DB:	10	Gaps: 23
 US-09-748-451-2 (1-338) x US-09-777-921A-3 (1-69327)		
OY	2	LeuGlnIysProIysSerValIyLsLeuArGlAlaLeuArgSerProIrgIysPheGluVal 21
Db	36404	ATTCAGAAAGCCTCAGAGCTG-----ATCGATTCAACTGGAAAGAGGTATC 36451
OY	22	Ala---GLyArg--SecCysglIngluValLeuArgIysGlyCysLeuArgPheGlnLeur 40
Db	36452	ACGCATTCAGAAATGAATGATGAATGAAGAAGCGAAGGAACTTTAGAAAAAAGAAAT 36511
OY	40	oGIuArGgLySerArg-----LeucY 47
Db	36512	AAAAGAGAAAGACAAAGCCCTCCAAATAATGAGACTGTGTGAAAGAACAAATCTATG 36571
OY	47	sleuTYr-----GluSprgLyThrGlnLeu 55
Db	36572	TCTGATTGGTGTAACCTGAAAGTACGAGGAGAAATGGAAACCAGTTGGAAACACTCTGCA 36631
OY	56	-----ThrluAsPrTyPheProSerValProAsPaAnaAlaGluValLeuLeuth 73
Db	36632	GGATATTATTCACAGAGAACTTCCCATGTGTAGCAAGSCAGGCCAAATTCAGATTCAGGA 36691
OY	73	rIeuglYglnAlATrPgInglYtyrValSerApriLeaGrpHeLeuSerAlaPheNI 93
Db	36692	AATPCACAGAAAGCCACAAAGATAC-----TCSTTGAAANAAGCAACTCCAAAGACAT 36745
OY	93	SGluPrOGlnIValIGlyLeuIlleglnAlaIagInglInleuLeuCysAsprGluAlAPr 113
Db	36746	AATTCACAGATTCCACCAAACTTAATGAAGAAAAATTTAAAGGCGCACAGAGAA 36805
OY	113	oGlnArGln-----ArgLeuLeuAlaerLeuLeuNizAsnVa 126
Db	36806	AGGTCCGGTACCCTCAATGGAAGCCATCAGACATAACAAGCGATCTCTTGCGCAAAAC 36865
OY	126	IserGlnAnIlaIalAgluThrArgIaGluAsrPorCoTrPrPheGlnGlyLeu-- 145
Db	36866	TCTACAA-----ACCAAGAGAGGTGGGGGCCAAATTTATACATTTCTTAA 36910
OY	146	-GluSerArgPheGlnSerIysSerGlyTy-----LeuArTyrSerCysGluSe 162
Db	36911	AGAAAAGAAATTTTCAACCCAGATTCATCATCCAGCCAAATAGGTTCATAAAGGAAG 36970
OY	162	rArgIleArGSetyr-----LeuArGlu---ValSerSetyrProse 176
Db	36971	AGAATTAATAATCTTTTACAGACAAAGCAAAATGCGTAGAAGATTTTGTCCACCACAGGCT-- 37028
OY	176	rThrValIGlyAlaGlnIuAgIngluPheLeuArgValLeuGlySerMetCysGlnAr 196
Db	37029	-----GCCCTAAAGAAAGTTCTGTGAAG-----GA 37051
OY	196	gleuArSerMetGlnTyAsnGlnSerTyPheAsparGlyAlaIySglyLySerAr 216
Db	37052	AGTGCCTTAACCTTGAAGAAACAAATCATGATAC-----CAGCGCGTCGCAAAATCATGCCAA 37105
OY	216	gleuCyStrPrroGluGlyTrPheSerCysGlnGlyProPheAspMetAsrSerCyle 236
Db	37106	AATGTAAAGACG-----TTCAGACT 37126

Score: 96.50
 Percent Similarity: 35.44%
 Best Local Similarity: 26.84%
 Query Match: 5.39%
 DB: 9
 Matches: 106
 Conservative: 34
 Mismatches: 125
 Indels: 131
 Gaps: 25

US-09-748-451-2 (1-338) x US-10-060-425-7 (1-2673)

6 LysSerValIysLeuArgAla---LeuArgSerProGlyLysPheGlyValAlaGlyArg 24
 2651 AAGAAGAAAGAGTGAAGGCGCACTTAGGGCAGCG----- 2616

25 SerCysGlnGluValLeuArgGlyGlyCysLeuArgPheGlnLeuProGlnArgGlySer 44
 2615 ---TGC---ACCGTGTACAGCGCAGTCTGTGATCTTTAC-GTGCCTCCGGGAGGGGA 2563

45 ArgLeuCys-----LeuArgLysAspGlyThrGlnLeu 55
 2562 CAGCTGTGTCTGCAATTGAGGAGCGAGTGGCTTGAAGCTGAGAGCGGGCAGCTTGGCT 2503

56 ThrGlnAspArgPheProSerValProAspAsnAlaGluLeuValLeuThrLeuGly 75
 2502 ACCGAG-----GGGGCCCTGAGAGATGTGTGACTGACTGATGAGGCTGCCCTGACG 2452

85 IleArgArgPheLeuSerAlaPheHisGlnProGlnValGlyLeuIleGlnAlaGln 104
 2391 GTGCTCTTCTTCGTGGCCGCGGT-----GCCGTGGTGGCGAAGGG 2350

105 GlnLeuLeuCysAspArgGlnAlaProGlnArgGlnArgLeuAlaAspLeuLeuHis 124
 2349 CATGCCACCTGTAT-----CTCAACTGTGAGGGGCTCAACTTCTT 2308

125 AsnValSerGlnAsnIleAla-----AlaGlnThr---ArgAlaGlnAspProGly 142
 2307 GATGTGGAGGGGTGCTGGCCAGCTCTTGAAGCAGAGAGCTCTCTCAGCGGG- 2249

142 HeuGlnGlyLeuGlnSerArgPheGlnSerLysSerGlyTyLeuArgTyrSerCysGln 162
 2248 -----ACGTGTACCAAGC-----TACAGATGGGTAGGCTC 2215

162 eArgGlnLeuSerTyLeuArgGlnValSerSerTyProSerThrValAlaGlyVal 180
 2214 ACCGTACAGCAGC-----GCATCAGCTACACCGAAGAGGAGCAGATGT 2167

181 ----GluAlaGlnGlnGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArg 199
 2166 GATGGCCGACTCGGCGC-----TGTGTGTCGATCTCGT 2134

199 eRmetGlnTyrAsnGlySerTyPheAspArg-----GlyAlaGlyGlySerA 216
 2133 CACTCGGACACTTGAAGCGGCTGTCCAGTACCTGTGGCTCCAGAGTGGC----- 2078

216 rGleuCysThrProGlnGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysL 236
 2077 -----TGCACAGATCTGGGCGCGGCGCATGTAGTCTCT----- 2042

236 euserArgHisSerIleAsnProTyrSerAsnArgGlnSerArgIleLeuPheSer---T 255
 2041 --TCCAGG-----CCCGTGGCCACACAGAGAACATATATGTCTCCAGGTGAG 1996

255 hrrTpaen-LeuAsp---HisIleIleGlnLysArgThrIleIleProThrLeuVal 273
 1995 TGTGGATTTGATACCTTTCATGCTTCCGA-----CCGGTACACGTA 1954

274 GluAlaIleLysGlnGlnAspGlyArgGlnValAspTyr---GluTyrPheTyrGly-Le 292
 1953 GAACACAGCAAAAGAGCAGATGGCGCGTGAAGCAGACAGAAATGAGCTTACTATAGAGCT 1894

292 u-----LeuPheThrSerGlnAsnLeuLysLeuValHis----- 303

1893 CCGAGTCAGGAGCTTGAACATTCACACACAGAAAGATGGCTGTGCCACAGCGGA 1834
 304 -----IleValCysHisLysLysLys 310

1833 AAGCAGGGGCGACGCTGCAGATCTGTGTGACCATGATCTTGTGAGGTCCAGTGAAG 1774
 310 rThrHisLysLeuAsnCysAspProSerArgIleTyrLysProGlnThrArgLeu----- 328

1773 GAACATCGGGGCAAACTGACAGTGGTCCCATGAGGCAAGCGCACACAGATGGGGAG 1714

329 -----LysArgLysGlnProValArgLysArg 337
 1713 GCGAAGAGAGAAAGAGAAAGTGGCCGATGAGCAGCAGCG 1677

RESULT 13
 US-09-995-898A-32
 ; Sequence 32, Application US/0995898A
 ; Publication No. US20030027253A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Presnell, Scott R.
 ; APPLICANT: Xu, Mengfeng
 ; APPLICANT: No. US20030027253A1ak, Julia E.
 ; APPLICANT: Whitmore, Theodore E.
 ; APPLICANT: Grant, Francis J.
 ; TITLE OF INVENTION: CYTOKINE RECEPTOR ZCYTOR19
 ; FILE REFERENCE: 00-108
 ; CURRENT APPLICATION NUMBER: US/09/995,898A
 ; PRIOR APPLICATION NUMBER: US 60/253,561
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: US 60/267,211
 ; PRIOR FILING DATE: 2001-02-07
 ; NUMBER OF SEQ ID NOS: 50
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 32
 ; LENGTH: 1922
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: MBP-human zcytor19 fusion protein polynucleotide
 ; OTHER INFORMATION: sequence
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (123)...(1922)
 ; US-09-995-898A-32

Alignment Scores:
 Pred. No.: 0.0503
 Score: 95.50
 Percent Similarity: 34.29%
 Best Local Similarity: 24.49%
 Query Match: 5.34%
 DB: 9
 Matches: 60
 Conservative: 24
 Mismatches: 101
 Indels: 60
 Gaps: 12

US-09-748-451-2 (1-338) x US-09-995-898A-32 (1-1922)

15 serProArgLysPheGlyValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCys 34
 1247 TCACACAGCGAATCGGTACCGCTGGTTCGCGCGTGG----- 1282

35 LeuArgPheGlnLeuProGlnArgGlySerArgLeuCysLeuTyrGlnAspGlyThrGln 54
 1283 -----ATCCAGGCGCGGTCTGGCCCTCCCGCAGAAATGT-----GACGCTGCT 1324

55 LeuThrGlnAspArgPheProSerValProAspAsnAlaGlnLeuValLeuLeuThrLeu 74
 1325 CTCCACAGAA-----CTTCAGCGGTGTACTGTGATGCTCC----- 1360

75 GlyGlnAlaArgGlnGlyTyrValSerAspIleArgArgPheLeuSerAlaPheHisGln 94
 1361 ----AGGGCTTGGAGCAACCCCAAGAGTGTACCTATTTTGTGGCTATGAGAGCTGTCCAC 1417

QY	30	Leu-----	ArgLeuGL1CyLeuAAspPheGlnLeuProGlyArg	42
Db	430	GTACACGACGACCCGATGCGACAGTAAAGTGCACATTCACCTGCAGATTCCAAACTTCCAAACCA	489	
QY	43	GlySerArgLeuGys-----	LeuTyGlnAAspGlyThrGlyLeu	55
Db	490	GGCATCTCAATTGTGTATGCGTGTCTCTGTGTACAAAGCTTTGTAAACCTGTGTGACGGAGTGT	549	
QY	56	ThrGlnAspTyGlyPheProSerValProAspAlaGlyLeu-----	Val170	
Db	550	CGTCCGCGCTATCCGCTATCCCGCTTGCACAAAGCCGACCTTTTGCAGAAAGATGTTCCCGCG	609	
QY	71	LeuLeuThrLeuGlyGlnAlaAlaTrpGlnGlyTyValSerAspIleArgArgPheLeuSer	90	
Db	610	GTGTGTGAAACCGCGCTGCGACAGCCGACGCGCTTT-----CGTTTTGTG-----	651	
QY	91	AlaPheHisGlnProGlnValGlyLeuIleGlnAlaAlaGlnGlnLeuLeuCysAspGln	110	
Db	652	GTTTTTCTGTGGAAACGATGGGAGATCTCTTCGTGTGCGACGACGCTGT-----	699	
QY	111	GlnAlaProGlnArgGlnArgLeuAlaAspLeuLeuHisAsnValSerGlnAsnIle	130	
Db	699	-----	699	
QY	131	AlaAlaGlnThrArgAlaGlnAspProPheTrpPheGlnGlyLeuGlnSerArgPheGln	150	
Db	700	CCCTCTGTTTCAAGCGCGGTGAAGTCCG-----GAACCGTTGACTTCACAG-----	744	
QY	151	SerIysSerGlyTyGlyLeuArg-----TyrSerCysGlnSerArgIleArgSerTyLeu	168	
Db	745	GCGCAGGAGCGAGTATCTGCTGTGTGATATACCTTCCGTGCGCGGAGGCGCTCA-----	798	
QY	169	ArgGlnValSerSerTyGlyProSerThrValGlyAlaGlnGlnGlnGlnPheLeuArg	188	
Db	799	-----GCGGAGTTTTCTGCGACGAGGATCTCGAAG	828	
QY	189	ValLeuGlySerMetCysGlnArg-----	LeuArgSerMetGlnTyr202	
Db	829	ACCTTCACGCGTACAGTGCACAGATCTTATGTGATGCGCATTCGCGCGCATTC-----	882	
QY	203	AsnGlySerTyPheAspArgGly-----	AlaIysGlyGlySerArgLeu217	
Db	883	-----TCTGTTTCCGTCGTGTGCGCAGTGTCCACAGACACCCTGTGATGAGACCCGCC	936	
QY	218	CysThrProGlnGlyTyGlyPheSerCysGlnGlyProPheAspMetAspSerCysLeuSer	237	
Db	937	ACGACCCCGGACGAGCGGTG-----CTCTCT	963	
QY	239	ArgHisSerIle-AsnProTySerAsn-----	ArgG1248	
Db	964	TCACCGCAGCTCAAAACCCGAGAAAGAACTCTCAGAAAGCTCGCGGAAAAACGAGGGA	1023	
QY	248	userArgIleLeuPheSerThrTrpAsnLeuAspHisIleIleGlnIysLysArgThr11	268	
Db	1024	AGAAAGAGCGCTGAAGTGCACACGAAAACTGAGTGTGTGCGTATGAGCAAGAGTGT	1083	
QY	268	eIleProThrLeuValGlnAlaIleIysGlnGlnAspGlyArgGlnValAspTrpGly	288	
Db	1084	CTTCCGACAGTGTGTC-----CGGATGTGTGTGGTATGCT	1119	
QY	288	PheTyGlyLeuLeuPheThrSerGlnAsnLeuIysLeuValHisIleValCysHisIy	308	
Db	1120	GTCAACGTGACTTCTGCTGTGTCGAAAGACCTGCTGCTGACCACTGCTGTCCATAAA	1179	
QY	308	SlyThrThrHisIysLeuAsnCysAspProSerArgIleTyGlyAspProGlnThrArg	328	
Db	1180	AAAGGTT-----GTTAACGCTCTTAAAAA	1209	
QY	328	uLysArgTyGln332		
Db	1210	AAAAAAAAAAAA1222		

Search completed: May 26, 2003, 16:58:10
Job time : 230 secs

```
OY 154 GlyTyrLeuArgTyrSerCysGluSerArgIleAlaArgSerTyrLeuArgGluValSerSer 173
Db 6102 ACTCGATTTGTTTCATGAGGACCACTAGTGTGCTCCACACAGGCCAGAGTGAGTCC 6043
OY 174 TyrProSerThrValGlyAlaGluAlaGluInlu---GluPheLeuArgValLeuGlySer 192
Db 6042 ---CCATCTACTGCTGGACGACAGGCGACCTTCTCTCCAGGCTTCACCCACTGGGGAGC 5986
OY 193 MetCysGlnArgLeuArgSerMetGlnTyrAsnGlySerTyrPheAspArgGlyAlaLys 212
Db 5985 CTTTGT-----ACATTCGAAAGTGAACCAAGTCATGGATGAA-----CAG 5944
OY 213 GlyGlySerArgLeuCysThrProGluGlyTyrP 223
Db 5943 GGTGGCCCGAGATTAGCATTCACAGAGGCTTG 5911

RESULT 20
US-09-764-891-9340/C
Sequence 9340, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 9340
LENGTH: 11598
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-9340

Alignment Scores:
Pred. No.: 4.35 Length: 11598
Score: 90.00 Matches: 36
Percent Similarity: 41.22% Conservative: 18
Best Local Similarity: 27.48% Mismatches: 63
Query Match: 5.03% Indels: 14
Db: 9 Gaps: 5

US-09-748-451-2 (1-338) x US-09-764-891-9340 (1-11598)
OY 94 GluProGlnValGlyLeuIleGlnAlaGluInluLeuCysAspGluGlnAlaPro 113
Db 6264 GAACACACAGGAGGCCCGGTACACACAGGATCTGTCTGTGTGCTCCCTAGCCAG 6205
OY 114 GlnArgGlnArgLeuAlaAspLeuIleAsnValSerGlnAsnIleAlaAlaGlu 133
Db 6204 CAGCTGGAGAACCCCTGGCC-----TGTGGGAACGTAGTCAGGGA 6163
OY 134 ThrArgAlaGluAspProTyrPheGluGlyLeuGluSerArgPheGlnSerLysSer 153
Db 6162 CCCATGTCACAGACAGAGGCCATTGGAGGAAATGGAAGCGATTAGCTCCAGTTAT 6103
OY 154 GlyTyrLeuArgTyrSerCysGluSerArgIleArgSerTyrLeuArgGluValSerSer 173
Db 6102 ACTCGATTTGTTTCATGAGGACCACTAGTGTGCTCCACACAGGCCAGAGTGAGTCC 6043
OY 174 TyrProSerThrValGlyAlaGluAlaGluInlu---GluPheLeuArgValLeuGlySer 192
Db 6042 ---CCATCTACTGCTGGACGACAGGCGACCTTCTCTCCAGGCTTCACCCACTGGGGAGC 5986
OY 193 MetCysGlnArgLeuArgSerMetGlnTyrAsnGlySerTyrPheAspArgGlyAlaLys 212
Db 5985 CTTTGT-----ACATTCGAAAGTGAACCAAGTCATGGATGAA-----CAG 5944
OY 213 GlyGlySerArgLeuCysThrProGluGlyTyrP 223
Db 5943 GGTGGCCCGAGATTAGCATTCACAGAGGCTTG 5911
```

Tue May 27 09:31:58 2003

us-09-748-451-2.rst

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK protein - nucleic search, using frame_plus_p2n model

Run on: May 26, 2003, 15:23:53 ; Search time 1707 Seconds

(Without alignments)
3206.839 Million cell updates/sec

Title: US-09-748-451-2

Sequence: 1 MCRKSVKRLRSPKRF.....SRVYKPTRLKRPVRRKQ 338

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seg length: 0

Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

*No. of hits
1000000*

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09748451/unat.21052003.153830.16291/app.query.fasta.1.519
-DB=EST -FASTA=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human0.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=20 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09748451.ecgn.1.1.1906.gunat.21052003.153830.16291 -NCP=6 -ICPU=3
-NO.MMAP -LARGEIOUERY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

EST: *
1: em_estha: *
2: em_esthum: *
3: em_estlin: *
4: em_estlun: *
5: em_estcov: *
6: em_estcpl: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_man: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	969	54.2	104	14	B0715165
2	957	53.5	796	13	B1657172
3	954.5	53.4	834	12	B0863734
4	871	48.7	685	13	B1517452
5	774	43.3	830	13	B1770511
6	761	42.5	855	12	B165535
7	738	41.3	775	13	B1916038
8	716	40.0	908	14	B0431184
9	664	37.1	1071	14	B0053507
10	638	35.7	689	9	AJ431881
11	588	32.9	716	12	B0705510
12	585	32.7	614	10	B0619078
13	576.5	32.2	744	9	AJ448138
14	576	32.0	637	9	AJ602837
15	572	30.8	526	17	AQ422321
16	551	30.2	515	12	BG086312
17	541	30.2	537	13	B1789500
18	518.5	29.0	643	10	B0628807
19	490	27.4	513	12	B0755721
20	487.5	27.2	688	13	BJ026619
21	475	26.6	446	10	B0849603
22	472	26.4	404	9	AL702256
23	449.5	25.1	540	13	BJ010671
24	444	24.8	567	13	BJ012356
25	438	24.5	1023	12	BG120825
26	401	22.4	689	12	BG190041
27	394.5	22.1	692	13	B1024478
28	385	21.5	713	12	BG194151
29	369	20.6	798	12	BG202840
30	352	19.7	581	13	BM128591
31	350	19.6	627	10	BH128945
32	347	19.4	719	12	BG194632
33	336	18.8	304	10	AM838362
34	332	18.6	851	10	BG139152
35	316	17.7	554	13	BM128641
36	288	16.7	473	9	AJ277912
37	273.5	15.3	358	13	AJ980778
38	273	15.3	358	10	AM487537
39	272.5	15.2	312	12	AM336839
40	272	15.2	210	13	BE936477
41	254.5	14.2	622	13	BG977108
42	254.5	14.2	802	9	AM489249
43	245	13.7	720	17	AJ394332
44	238	13.3	766	13	AJ284961
45					B1857018

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
B0715165	LOCUS	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	DEFINITION	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	ACCESSION	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	VERSION	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	KEYWORDS	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	SOURCE	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	ORGANISM	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	REFERENCE	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	AUTHORS	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	
B0715165	TITLE	AGENCOCURT_8480637 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310003	B0715165	5/	EST	house mouse.	Mus musculus	1014 bp	EST 16-JUN-2002	

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: Resgen, Invitrogen Corp.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM13731 row: j column: 20
High quality sequence stop: 566.
Location/Qualifiers

FEATURES
Source

1. 1014
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6310003"
/clone_1lb="NIH-MGC.129"
/lab_host="DH10B (phage-resistant)"
/note="Organ: olfactory epithelium; Vector: pcwv-sport6.1.cdb; Site_1: EcoRV; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.2 kb. Constructed by Resgen, Invitrogen Corp. Note: this is a NIH-MGC library."
BASE COUNT 238 a 266 c 303 g 207 t

Alignment Scores:

Pred. NO.: 3.73e-105 Length: 1014
Score: 969.00 Matches: 176
Percent Similarity: 89.95% Conservative: 21
Best Local Similarity: 80.37% Mismatches: 20
Query Match: 54.16% Indels: 2
DB: 14 Gaps: 1

US-09-748-451-2 (1-338) x BQ751565 (1-1014)

OY 120 AlaspleuenuhtsAnValserGlnAsnIleAlaIaGluThrArgAlaGluAspPro 139

DB 1 GCCGATCTTCGTGATCAGCGACCCAGAAATTAATCTGACAGACCCGCGAGACGACCA 60

OY 140 ProTrpPheGluGluGluSerArgPheGluSerGlyTyrLeuArgTyrSer 159

DB 61 TCCTGTTTAAAGCTTTGGAGTCGAGATTAGGAATTAATCGGGCTATCTGAGTACAC 120

OY 160 CysGluSerArgIleArgSerTyrLeuArgGluValSerTyrProSerThrValGly 179

DB 121 TGTGAGAGTCGAGTCCGGGGTACCTTAAGAGAGGTAGAGCGCTTACCTTATGTCGAT 180

OY 180 AlaGluAlaGluGluGluPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSer 199

DB 181 GAAGCAGCTTCAGAAAGAGTACGCTGCGGCTCTGCTCCATGTCGCAAGCTCAATCG 240

OY 200 MetGlnTyrAsnGlySerTyrPheAspArgIleValAlaGlyGlySerArgLeuGlyThr 219

DB 241 GTGCAGTACAAATGCGACCTATTCGACAGAGGTGACAGAGCCAGCCGCTCTGACT 300

OY 220 ProGluGlyTyrPheSerCysGlnGlyTyrPheAspMetAspSerCysLeuSerArgHis 239

DB 301 CCAGAGAGTATGTTCTCTCCAGAGGCCCTTGTACCTGAGAGCTCTTCCCAACAC 360

OY 240 SerIleAsnProTyrSerAsnArgGluSerArgIleLeuPheSerThrTrpAsnLeuAsp 259

DB 361 TCCATCAACCCCTATGCGACAGAGAGCCGAGTCTCTTCAGTACCTGGAACCTGGAT 420

OY 260 HisIleIleGluIleGlySerArgThrIleIleProThrLeuValGluAlaIleGluGln 279

DB 421 CATATATATAGAGAGAGCCACCGGTGATCCACGCTGCTGAAGCCATC-----CAG 474

OY 280 AspGlyArgGluValAspTrpGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLeu 299

DB 475 GATGGGAGGAGGTGAAGTGGAGTACTTCTACAGCTGCTCTTCACTGCGAGAACCTG 534

OY 300 LysLeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSer 319

DB 535 AAGTGTGTGACATCGCTGCTCCACAGAACACACACAGAGGTGAGTCCGACCGCAGT 594

OY 320 ArgIleTyrAspProGlnTrpArgLeuArgGlyGlnProValArgLysArgGln 338

DB 595 AGGATCTATGCGCTCAGACAGATCCAGAGGAGGAGGAGGCTGCTCGAAGAGCGC 651

RESULT 2 796 bp mRNA linear EST 12-SEP-2001

LOCUS 603280838F1 NCL.CGAP_Mam4 Mus musculus cDNA clone IMAGE:5325440-3

DEFINITION mRNA sequence.

ACCESSION BI657172

VERSION BI657172.1 GI:15571408

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Lotmar Hennighausen Ph.D., Priscilla Furth Ph.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:

http://image.llnl.gov

Plate: LLM1825 row: k column: 09

High quality sequence stop: 762.

Location/Qualifiers

1. 796

/organism="Mus musculus"

/strain="NMRI"

/db_xref="taxon:10090"

/clone="IMAGE:5325440"

/clone_1lb="NCL.CGAP_Mam4"

/tissue_type="tumor, gross tissue"

/dev_stage="5 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pcwv-sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lotmar Hennighausen/Priscilla Furth, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."

BASE COUNT 172 a 226 c 235 g 163 t

ORIGIN

US-09-748-451-2 (1-338) x BI657172 (1-796)

OY 1 MetLeuGlnLysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20

DB 55 GTGCTCGCCCAACCCAAATGCTGAGAGTCCATATGCGCTGCAATGCTGCG 114

OY 21 ValAlaGlyArgSerCysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuPro 40

DB 115 GTGGCGCGCCGAGAGCTCCAGAGAGCTGCTGTAAGGCTGCGTCCGCTTCAAGCTCCG 174

Alignment Scores:

Pred. NO.: 7.01e-104 Length: 796
Score: 957.00 Matches: 185
Percent Similarity: 86.29% Conservative: 29
Best Local Similarity: 74.60% Mismatches: 34
Query Match: 53.49% Indels: 2
DB: 13 Gaps: 0

OY		41	GluA9gylSerATgLeucySLeuTYrGluaspLyTrGluLeuPThrgLysAspTYrPe	60
Db		175	AMGCCCGGTTCGGAGTGCCTGACATGAAGAATGGCAGGAGGACGAGCACTGCTTC	234
OY		61	ProSeArVtIProApaspaNAlagIueuValleuLeuPThreugLYglNAlarTPngInly	80
Db		235	CCGGGCGCTTCOCMAACACAGCTAGACTCCATTGCTGCACCGCTGGCAGACCTGGCATGGC	294
OY		81	ThyrValSerAspIleArGrPhelLeuSerAlArpheNhgIuproGlnValAglyLeuile	100
Db		255	TATGTGATGATCACATCACACTTTCTTCAGTGCTTTTAATGACCACATGCGCGGCTATC	354
OY		101	GlnAlalagIngleuLeuCySaSpagIunAlaprogInArGlnArGleuLeuala	120
Db		355	CAGGCTGCACGGCACTGCTGTGATGTAGCAGGGCCCCACCTGAGGCAAAGAGTGGTGGGCC	414
OY		121	AspleuLeuHIsaenValSerGlnasnIlleaIalagIuthraArgalaGuaspPropo	140
Db		415	GATCTCTGCATCTCAGTGCAGACCGAAATATTACTGACAGACCCGGGAGACAGA-CCATCC	473
OY		141	TrpPheglugIleuGlusEArGrPheglInserTysserGlyTYrLeuArgrTyserCyS	160
Db		477	TGCTTTGAAAGGTTTGGAGTGCAGATTTCACAGATAAGTGGGCTATCTGAGATACACTCT	533
OY		161	GluseArGrIleArSeserTYrLeuArGrGluValSerseTYrProserThrValglyAla	180
Db		534	GAGACTGCGATCCGGGGTTCCTTAAGAGGGTGCAGACCGTTCACACTCTATAGGGATGAA	593
OY		181	GluAlagIngIugIuphelLeuArGrValleuGlyserMetCYsGlnArGleuArGrserMt	200
Db		594	GCAGCTCAAGAAAGATACCTGCGAGTCTTGCTGCATGTCGAGAAAGCTCAAAATGGGtg	653
OY		201	GlnTYrAnsgIysErTYrPheAspArGrGIValAlaySGIyLSerArGleuCyGrmpPro	220
Db		654	CAGTCAATATGGCACACTTATTTCAGCAGAGGTGCAGAACCGCACGCCGCTC-TGTACTCCA	712
OY		221	GlnGlyTYrPheSerCySgInGlyArProPheAspMeTapSerCYsLeuSerArghISer	240
Db		713	GAAAGATGCTTCTCTCCGCCAGGGCCCCCTTTGACTGAGAGACTGTCTTTCCAAAGCACTCC	772
OY		241	IleAsnProrTYrSerAsnArGrIu	248
Db		773	ATCAAAACCTATGGCCACAGCAGAG	796
RESULT 3	BG863734			
SEQUENCE	602979671761 NCI CGAP Mam4 Mus musculus CDNA clone IMAGE:431772 5'	834 bp mRNA linear	EST 29-MAY-2001	
MISSION	MRNA sequence.			
VERSION	BG863734			
KEYWORDS	BG863734.1 GI:14214272			
SOURCE	EST.			
ORGANISM	house mouse.			
	Mus musculus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
	1 (bases 1 to 834)			
REFERENCE	NIH-MGC http://mgc.nci.nih.gov/			
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)			
TITLE	Unpublished (1999)			
JOURNAL	Contact: Robert Straube, Ph.D.			
COMMENT	Email: csepbes-remail.nih.gov Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch Ph.D. cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Cloned distribution: MGC clone distribution Information can be found through the I.M.A.G.E. Consortium/LNL at: http://Image.lnl.gov Plate: LHAM10830 row: a column: 05 High quality sequence stop: 788. Location/Qualifiers			

```

source
1. 834
/organism="Mus musculus"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:4917772"
/clone_lib="NCI-GCAP_MamA"
/tissue_type="tumor.gross tissue"
/dev_stage="45 months"
/lab_host="DH10B"
notes: Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt; Site: 2; Motif: Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Jøhan Henningsen/Priscilla Fuths, NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1996)."
BASE COUNT      186 a      236 c      240 g      172 t
ORIGIN
Alignment Scores:
Pred. No.:      1,51e-103      Length:      834
Score:          954.50      Matches:      190
Percent Similarity: 83.59%      Conservative: 29
Best Local Similarity: 72.52%      Mismatches: 39
Query Match:      53.35%      Indels:      5
DB:              12      Gaps:      1
US-09-748-451-2 (1-338) x BG863734 (1-834)
Oy      1 MetLeuGlnIysProIysSerValIysLeuArgAlaLeuArgSerProArgIysPheGly 20
Db      47 GGGCTCGCCCAACCCAAATGCTGACATGTCGAGCCCTCAATAGCGCTGCACAGTTGGC 106
Oy      21 ValAlaGlyArgSerCysGlnGlnValAlaLeuArgIysGlyCysLeuArgPheGlnLeuPro 40
Db      107 GTGGCGGGCCGGAGCTCCGAGAGCTGCTGCTGTAAGGCTGGCTGCTTCACACTCCG 166
Oy      41 GlnArgGlySerArgIleuGlyLeuGlyGlnAspGlyThrGlnLeuThrGlnAspTyrPhe 60
Db      167 ATGCCCGGTTCCGGCTGTGCTGTGACGAAATGCGACGGAGTACGACGAGCTGCTTC 226
Oy      61 ProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaATPGLngly 80
Db      227 CCGGGCCCTCCCAACGACGCTGAGTCTCTATATCTGCTACCGCTGGGAGACCTGACATGGC 286
Oy      81 TyrValSerAspIleArgTyrPheLeuSerAlaPheHisGluProGlnValGlyLeuIle 100
Db      287 TATGTGAGTACATCCACACAGTTTCTCAGTGTGTTTAAAGACCCACATGCCGCTCATC 346
Oy      101 GlnAlaAlaGlnGlnIleuLeuGlyCysAspGlnGlnAlaProGlnArgGlnArgLeuAla 120
Db      347 CAGGCTGCACGGCAACATGCTGTGCAAGTAGACAGGCCCCACTAGCGCAAAAGCTGCGCC 406
Oy      121 AspLeuLeuHisAsnValSerGlnAsnGlnAlaAlaGlnThrArgAlaGluAspPropo 140
Db      407 GATCTCTGCATCAGCTGAGCCAGAAATTAATTAATGTCAGAGAACCCGGAGAGCAAGCATGC 466
Oy      141 ThrPheGlnGlyLeuGlnIleuSerArgPheGlnIleuSerIysSerGlyTyrLeuArgTyrSer 160
Db      467 TGGTTTGAAGGTTAGAGAGTCGAGATCCAGAAATTAATTCGGCTATCTAGATACACTGT 526
Oy      161 GluSerArgIleArgSerTyrIleuArgGlnValSerSerTyrProSerThrValGlyAla 180
Db      527 GAGAGTGTGCAATCCGGGGCTTACTTAAGAGAGGTACAGCGCTTAACCTCTATGTGTGAGAA 586
Oy      181 GluAlaGlnGlnIleuPheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMet 200
Db      587 GCAGCTCAAAAGAGTACTGCGAATCTTGGCTCCATGTGTGCAAAAGCTCAAAATGGGG 646
Oy      201 GlnTyrAsnGlySerTyrPheAspArgGlyAla---LysGlyGlySerArgLeuGlyCysThr 219
Db      647 CAGTACAAATGACAGCTATTATTCACACGAAGGTGCGAAGAACCCGACGACGCTCTACT 706
Oy      220 ProGlu-GlyTyrPheSerCysGlnGlyProPhe-AspMetAspSerCysLeuSerArgH 239

```

```

Db      707 CCAGACGATGTTCTCTCCAGAGGCCCTTGGACCTGAGAGCTGCTTCCAGC 766
QY      239 ISSerIleAsnProTySerAsnArg-GluSerArgIleuPheSerThrTriPAsnIeu 258
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      767 ACTCAATCAGACCTTATGCAAGAACAGAGGAGATCTCTCAGTCC-TGAGACCTG 825

RESULT 4
BIS17452
LOCUS   685 bp mRNA linear EST 29-AUG-2001
DEFINITION 603041725F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5162791 5'
ACCESSION B1517452
VERSION   B1517452
KEYWORDS  GI:15342244
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 685)
           NIH-MGC http://mhc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
           DNA sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LMNL at:
           http://image.llnl.gov
           Plate: LHAM11404 row: b column: 08
           High quality sequence start: 6
           High quality sequence stop: 685.
FEATURES
     source
         1..685
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5162791"
            /clone_id="NIH_MGC_116"
            /lab_host="DH10B"
            /note="Organ: pooled colon, kidney, stomach; Vector:
            PCWV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
            source anonymous pool of 3 colons, age 26 yo male, 49 yo
            female, 71 yo male colon; 46 yo male kidney, and pool of 2
            stomachs, 62 yo male and 70 yo female. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.4 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            023. Note: this is a NIH-MGC library."
COUNT  121 a 202 c 236 g 126 t
ALIGN
Alignment Scores:
Pred. No.: 1.15e-93 Length: 685
Score: 871.00 Matches: 179
Percent Similarity: 89.95% Conservative: 0
Best local Similarity: 89.95% Mismatches: 2
Query Match: 48.69% Indels: 19
DB: 13 Gaps: 1
US-09-748-451-2 (1-338) x BIS17452 (1-685)
QY      1 MetLeuGlnIysProLysSerValLysLeuArgAlaLeuArgSerProArgLysPheGly 20
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      91 ATGCTCCAGAGGCCCAAGAGCGTGAAGCTGCGGCGCTCGCACCCGAGGAAATTCGGC 150
QY      21 ValAlaGlyArgSerCysGlnGluValLeuArg-LysGlyCysLeuArgPheGlnIeuPr 40
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      151 GTGCTGCGCGGAGACTCTCCAGAGAGCTGCTGTCGAAGGCGTGTCTCCGCTTCCAGCTCC 210

```

```

QY      40 OGluArgGlySerArgLeuCysLeuTyrgLysPoiYThrGluLeuThrGluAspTyrg 60
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      211 TGACGCGCGTTCGCCGCTGTGCTGTACGAGGATGCGACGAGAGCTGACGGAATTTACTT 270
QY      60 eProSerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaTrpGlnI 80
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      271 CCCAGTGTCCCGCAACAGCCGAGCTGTCTCTCCTCCTGAGCCAGGCGCTGGCAGCG 330
QY      80 YTyrr----- -Valse 83
        |||||
Db      331 CTG-ATGTCCTTCTGCTGAGACCGGACCTTTTGTGTCTCCATTGGTCGAGATGTGAG 389
QY      83 rAspIleArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAlaI 103
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      390 CGACATCAGGCGCTTCTCCTGATGATTCACAGAGCAGACAGTGGGGCTCATCCAGCGCGC 449
QY      103 agIinGlnLeuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuLeuAlaAspLeu 123
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      450 CCACAGCTGTGTGTGTGTGAGCAGGCGCCACAGAGGAGGAGGCTGCTGCTGACCTTCT 509
QY      123 uHisAsnValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspPropTorpPheG 143
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      510 GCACAGCTCAGCCAGAACATCGCGCGGAGACCCGGCTAGACACCGCGCTGTTGA 569
QY      143 uGlyLeuGluSerArgPheGlnSerLysSerGlyTyrlleuArgTySerCysGluSerAr 163
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      570 AGGCTTGAGTCCGATTCAGACAGTGTGCTATCTGATGATACAGCTGTGAGAGCGC 629
QY      163 gIlaArgSerTyrlleuArgGluValSerSerTyProSerThrValGlyAlaGlu 181
        |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      630 GATCCGAGTTACTGAGGAGGTGAGTCTTACCCCTCCAGCGGTGGGTGGCGAG 684

RESULT 5
B1770511
LOCUS   830 bp mRNA linear EST 25-SEP-2001
DEFINITION 603055488F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5204996 5'
ACCESSION B1770511
VERSION   B1770511.1 GI:15762089
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 830)
           NIH-MGC http://mhc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
           Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-remail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
           DNA sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LMNL at:
           http://image.llnl.gov
           Plate: LHAM11513 row: p column: 21
           High quality sequence stop: 820.
FEATURES
     source
         1..830
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:5204996"
            /clone_id="NIH_MGC_122"
            /lab_host="DH10B"
            /note="Organ: pooled lung and spleen; Vector: PCWV-SPORT6;
            Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
            anonymous pool of 24 week female lung, 16 week female
            spleen, and 20-22 week male spleens. Library is oligo-dT
            primed and directionally cloned (EcoRV site is destroyed
            upon cloning). Average insert size 1.4 kb, insert size

```


OY 200 tGtTyrAsnGlySerTyrPheAspArgGlyAlaIysGlySerArgLeuCysThr-- 219
 Db 660 GCACCTACAAATGAGGATATTCGACAGAGGTGGGAAAGCCAGCCCTCTGTCACG 719
 OY 220 ProGluGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCys 235
 Db 720 AAGGAGGAT-----TCCTGGCAGGGGCTTGTACCTGGGAGGGGTGT 761

RESULT 7
 LOCUS B1916035 775 bp mRNA linear EST 16-OCT-2001
 DEFINITION 603184463P1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248694 5'
 mRNA sequence.
 ACCESSION B1916035
 VERSION B1916035.1 GI:16179992
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 775)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: L1AM11627 row: m column: 15
 High quality sequence stop: 709.
 Location/Qualifiers
 1. 775
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5248694"
 /clone_11d="NIH_MGC_121"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
 fetal brains, female age 20 weeks, female age 24 weeks,
 and male age 26 weeks. Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 0.7-3.5 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 017.. Note:
 this is a NIH_MGC Library."

FEATURES

Alignment Scores:
 Pred. No: 1.28e-77 Length: 775
 Score: 738.00 Matches: 162
 Percent Similarity: 88.00% Conservative: 14
 Best local Similarity: 81.00% Mismatches: 20
 Query Match: 41.25% Indels: 6
 DB: 13 Gaps: 0

US-09-748-451-2 (1-338) x B1916035 (1-775)

COUNT 151 a 233 c 239 g 152 t

OY 62 SerValProAspAsnAlaGluLeuValLeuLeuThrLeuGlyGlnAlaTrrP-GlnGlyTyr 81
 Db 740 AGCGTCTCTGACGATCCCGAGCTGGTCTGCTCAGCCAGCGCTGGCCAGGCTA 681
 OY 81 ValSerAspIleArgPheLeuSerAlaPheHisGluProGlnAlaGlyLeuIleG1 101
 Db 680 TGTGAGTGAACAT-CGGTGGCTTCCCACTGGCTTTCCGACAGCGCAGCGTGGCTCATCA 622

OY 101 n-AlaAlaGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 Db 621 AGCGCCCGCCGACACTACTGTGATGATGACAGGCCCCACAGAGGACAGACTCTGGCC 562
 OY 121 AspLeuLeuHisAsnValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspProPro 140
 Db 561 GACCTCTGTGAAACGTCAGCCAGACACCTCCAGCCGACCCGCGGCTGAGAGACCCCTG 502
 OY 141 TrpPheGluGlyLeuGluSerArgPheGlnSerIleLeuLeuLeuLeuLeuLeuLeuLeu 160
 Db 501 TGTGTTGAGGCTTGAGTCCGATTTGGAAATAGTGTGCTATCTGAGATACAGCTGT 442
 OY 161 GluSerArgIleArgSerTyrLeuArgGluValSerSerTyrProSerThrValGlyAla 180
 Db 441 GAAAGCCGAGATCCGAGGTACCTAGAGAGGTGAGCTCTCCATGCTGCTGGTGCA 382
 OY 181 GluAlaGlnGluGluPheLeuArgValLeuGlySerMetCysGlnAlaGluLeuArgSerMet 200
 Db 381 GAGCTCAGAGGAAATTCCTGCGGGGCTCTCCGTCATGTCAGAAAGCTCCACTGTGTG 322
 OY 201 GlnTyrAsnGlySerTyrPheAspArgGlyAlaIysGlySerArgLeuCysThrPro 220
 Db 321 CAGTACAGAGCGGACACTATAC-AACAGAGAGCCAAAGGTGGACCCGCTCTGACACCG 263
 OY 221 GluGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSer 240
 Db 262 GAAAGCTAGTCTCTCTCCGACAGGCTCCCTTCACATGACGAGGCTGTCAATATCTCC 203
 OY 241 IleAsnProTyrSerAsn-ArgGluSerArgIleLeuPheSerThrTrpAsn 257
 Db 202 ATCAACCCCTACAGTACACGAGAGAGGCTCTCTCTGAGACTGTGAAC 151

RESULT 8
 LOCUS BQ431184 908 bp mRNA linear EST 24-MAY-2002
 DEFINITION AGNCOURT_7786920 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5088786
 5', mRNA sequence.
 ACCESSION BQ431184
 VERSION BQ431184.1 GI:21170260
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 908)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: L1AM13350 row: h column: 03
 High quality sequence stop: 595.
 Location/Qualifiers
 1. 908
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5088786"
 /clone_11d="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

FEATURES

Alignment Scores:
 Pred. No: 1.28e-77 Length: 775
 Score: 738.00 Matches: 162
 Percent Similarity: 88.00% Conservative: 14
 Best local Similarity: 81.00% Mismatches: 20
 Query Match: 41.25% Indels: 6
 DB: 13 Gaps: 0

US-09-748-451-2 (1-338) x B1916035 (1-775)

COUNT 151 a 233 c 239 g 152 t

```

BASE COUNT      158 a      271 c      307 g      171 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      7.15e-75      Length:      908
Score:          716.00      Matches:      169
Percent Similarity: 87.318      Conservative: 3
Best Local Similarity: 85.798      Mismatches: 13
Query Match:    40.028      Indels:      12
DB:              14      Gaps:          2

US-09-748-451-2 (1-338) x BQ431184 (1-908)

QY      1 MetLeuGlnIysProIysSerValIysLeuArgAlaLeuArgSerProArgIysPheGly 20
        |||||||
        312 ATGCTCCAGAGAGCCCAAGAGCGTGAAGCTGGCGGCGCTCGCGAGCCGAGGAGGAGTGGCG 371
        |||||||
        21 ValAlaGlyArgSerGlyGlnGlnValLeuArgIysGlyCysLeuArgPheGlnLeuPro 40
        |||||||
        372 GTGGCTGGCGGAGAGTCCAGAGAGTGTGGCGAAGGGGCTGTCTCGGCTTCCAGGCTCCCT 431
        |||||||
        41 GluArgIysSerArgIysLeuArgIysLeuArgIysLeuArgIysLeuArgIysLeuArg 60
        |||||||
        432 GAGCGCGGCTTCCGCGCTGTGCTGTACAGAGTGGCGAGAGCTGAGCGAGAGTATCTTC 491
        |||||||
        61 ProSerValProAspAsnAlaGlnLeuValLeuLeuThrLeuGlnValGlnAlaArgIys 80
        |||||||
        492 CCGAGGTCTCCGACAAAGCGAGCGTGGCTGTCTGAGTTCATTTTCAGAGCCGAGGCTGGC 551
        |||||||
        81 TyrValSerAspIleArgArgPheLeuSerAla-PheHisGlnProGlnValGlyLeu 100
        |||||||
        552 TATGTGAGCAGACATCAGGCGCTCTCTCAGTTCATTTTCAGAGCCGAGGCTGGC 611
        |||||||
        100 eGlnAla-AlaGlnGlnLeuLeuCysAspGlnGlnAlaProGlnArgGln-ArgLeuLeu 119
        |||||||
        612 CCAGCGCCCGCCAGCAGCTGCTGTGTGATGAGAGCCCGCCAGAGAGAGAGAGGCTGCTG 671
        |||||||
        120 AlaAspLeuLeuHisAsnValSerGlnAsnIleAla-AlaGlnThrArgAlaGlnAsp 139
        |||||||
        672 GCTGACCTCTCCGACAGACGTCAGCCAGAAACATCCGCGGCGAGAGCCGCGGCTGAGACCC 731
        |||||||
        139 oProTIP-PheGlnGlyLeuGlu---SerArgPheGlnSerIysSer-GlyTyrLeu-Arg 157
        |||||||
        732 CGCGTGGTTTGAAGGGCTGGAAATCCGAAATTCAGAGAGAGTCTGGCTATCTGAGAG 791
        |||||||
        157 gTyr-SerCysGlnSerArgIysArgIysSer-TyrLeuArg-GluValSerSerTyr---Pr 175
        |||||||
        792 ATACACTGTGAAAGCCGAGATCCGAGATTACCTGAGAGGAGAGTACGTTCTTACCCCC 851
        |||||||
        175 cSerThrValGlyAlaGlnAlaGlnGlu 185
        |||||||
        852 TTCACAGGTGGGTGCCGAGAGCCCTCCAGAA 882

RESULT 9
BQ055307/c      1071 bp      mRNA      linear      [EST 29-MAR-2002]
LOCUS          AGENCOURT.6797380 NIH_MGC_99 Homo sapiens cDNA clone IMAGE:5807463
DEFINITION    5', mRNA sequence.
ACCESSION    BQ055307
VERSION      BQ055307.1 GI:19814647
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS      NIH-MGC http://mgi.cni.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgaaps-remail.nih.gov
              Tissue Procurement: Lou Staudt
              cDNA Library Preparation: Rubin Laboratory

```

```

BASE COUNT      248 a      344 c      278 g      201 t
ORIGIN

Alignment Scores:
Pred. No.:      1.61e-68      Length:      1071
Score:          664.00      Matches:      145
Percent Similarity: 83.77%      Conservative: 15
Best Local Similarity: 75.92%      Mismatches: 19
Query Match:    37.12%      Indels:      13
DB:              14      Gaps:          2

US-09-748-451-2 (1-338) x BQ055307 (1-1071)

QY      76 GlnAlaTrpGlnGlyTyrVal-----SerAspIleArgArgPheLeuSer 90
        |||||||
        921 AGGCGTGGCGAGGCGCTTATGAGTACACATCCGCGGCGCTTCTCAGGTGCTT----- 868
        |||||||
        91 AlaPheHisGlnProGlnValGlyLeuIleGlnAla-AlaGlnGlnLeuLeuCysAsp 110
        |||||||
        867 ---TTCGCAAGCGGAGCGGCGGCTTATCCAGAGCCGCGGCGAGTACTGTGTATGA 811
        |||||||
        110 uGlnAla-ProGlnArgGlnArgIysLeuLeuAlaAspLeuLeuHisAsnValSerGlnAsn 130
        |||||||
        810 GCAGGCCCGCCAGAGGAGAGACATCTCTGCCAGCTCTCTGTACAGTACAGCAAGAG 751
        |||||||
        130 le-AlaAlaGlnThrArgAlaGlnAspProPro-TripheGlnGlyLeuLeuSerArg 149
        |||||||
        750 TCCGAGCGCCAGACCGCGGCTGAGACCCCGCTGTGTGAAGCTTGAAGTCCCGATT 691
        |||||||
        149 eGlnSerIysSerGlyTyrLeuArgTyrSerGlnSerArgIleArgSerTyrLeuArg 169
        |||||||
        690 TTGGAAATAGCTGGCTATCTGTGATACAGCTGTGAAGCCGAGATCCGAGTACTCTGAG 631
        |||||||
        169 gGluValSerSerTyrProSerThrValGlyAlaGlnAlaGlnGlnGluValPheLeuArg 189
        |||||||
        630 AGAGGTGAGCTCTGCTCTCTCATGGTGTGGTGCAGAGGCTGAGAGGAATTCCTCGCGGT 571
        |||||||
        189 IleuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsnGlySerTyrPheAsp 209
        |||||||
        570 CTCCTCGCTCCATGTGGCCAGAGAGCTCAGCTGTGAGTACAGCGCAGCTATAC-AAAG 512
        |||||||
        209 gGlyAlaValGlyIysSerArgIysLeuArgIysThrProGlnGlyTyrPheSerCysGln 229
        |||||||
        511 AGAGCCCAAGAGGTGGAGCGGCTGTCTGCAACCCGGAAGCAGATTCCTCGCCAGGCTCC 452
        |||||||
        229 oPheAspMetAspSerCysIysLeuSerArgIysSerIleAsnProIysSerAsnArgGlu 249
        |||||||
        451 CTTCACATGAGACGCTGCTTGTCAAGATACATCCATCAACCCCTACAGTACAGGAGAG 392

```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LUCM2049 row: 0 column: 16
 High quality sequence stop: 595.
 Location/Qualifiers

FEATURES

source

1. 1071
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5807463"
 /clone_lib="NIH_MGC_99"
 /issue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Lymph. Vector: pORF7; Site.1: XhoI; Site.2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GGCACGAG(G). Size-selected >500bp for average insert size
 1.8kb. Library constructed by Ling Hong in the laboratory
 of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

OY 249 TArglleuPheSerThrTTPasn 257
 DB 391 CAGGCTCCTCTTGGACCTGGAC 367
 RESULT 10
 LOCUS AJ451681 689 bp mRNA linear EST 22-APR-2002
 DEFINITION AJ451681 riken1 Gallus gallus cDNA clone 29j11r1, mRNA sequence.
 ACCESSION AJ451681
 VERSION AJ451681.1 GI:20261777
 KEYWORDS EST.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 689)
 Buerstedde, J. M.
 Gallus gallus bursal lymphocyte EST
 Unpublished (2002)
 CONTACT: Buerstedde JM
 Cellular Immunology
 Heinrich-Pette-Institute
 Martinstr. 52, 20251 Hamburg
 Email: URL: http://genetics.mp1.uni-hamburg.de/dta0est.html.
 Location/Qualifiers
 1..689
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="29j11r1"
 /clone_11b="8riken1"
 /cell_type="bursal lymphocyte"
 /dev_stage="2-3 weeks old"
 /note="CB inbred strain"
 BASE COUNT 166 a 172 c 219 g 132 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.06e-65 Length: 689
 Score: 638.00 Matches: 130
 Percent Similarity: 51.60% Conservative: 31
 Best Local Similarity: 41.67% Mismatches: 48
 Query Match: 35.66% Indels: 103
 DB: 9 Gaps: 2
 US-09-748-451-2 (1-338) x AJ451681 (1-689)
 OY 6 LysSerValLysLeuArgAlaLeuArgSerProArgLysPheGlyValAlaGlyArgSer 25
 DB 56 AGGGGTTCCCGCTCCGCCCGCCGCGCAGCGCGCAAGTTTCGGGGCGCGCGCGGAGC 115
 26 CysGlnGluValLeuArgLysGlyCysLeuArgPheGlnLeuProGluArgGlySerArg 45
 116 CTGCGTGGCGTGGCGGAGGCGGTGGCGGTGCTGCAAGCTCCGTTGGCAGCGCGCGG 175
 OY 46 LeuCysLeuTyrGluAspGlyThrGluLeuThrGluAspTyrTrpPheProSerValProAsp 65
 DB 176 CTGTGCTCTACGAGGAGCGGAGCGGAGCTGCTCCGACGCTGCGCGCGG 235
 OY 66 AsnAlaGluValLeuLeuThrLeuGlnAlaTrpGlnGlyTyrValSerAspIle 85
 DB 236 CAGACGGAGCTGGCTGCTGCGCGCGCGGAGAGCTGCGCGCGG 280
 OY 86 ArgArgPheLeuSerAlaPheHisGluProGlnValGlyLeuIleGlnAlaAlaGlnGln 105
 DB 280 280
 OY 106 LeuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuAlaAspLeuLeuHisAsn 125
 DB 280 280
 OY 126 ValSerGlnAsnIleAlaAlaGluThrArgAlaGluAspProProTrpPheGluGluLeu 145

DB 280
 OY 146 GluSerArgPheGlnSerLysSerGlyTyrLeuArgTyrSerCysGluSerArgIleArg 165
 DB 280
 OY 166 SerTyrLeuArgGluValSerSerTyrProSerThrValGlyAlaGlnAlaGlnGluGln 185
 DB 281
 OY 186 PheLeuArgValLeuGlySerMetCysGlnArgLeuArgSerMetGlnTyrAsnGlySer 205
 DB 293 TATAAAGGATATCTGACCTGATGTTGATTAACGAAATCTGTGAAGATTAACGCGACG 352
 OY 206 TyrPheAspArg---GlyAlaLysGlyLysSerArgLeuCysThrProGluGlyTrpPhe 224
 DB 353 TACTTTGACAGCAGACAGAGAGAGCGCGGAGATGCCCTGTGCACAAAGAGGAGATGTTTC 412
 OY 225 SerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHisSerIleAsnProTyr 244
 DB 413 TCTTGTACAGGGCCCTTTTGACAGAGATGATCCCATGTATGATTAACATTAACCCCTAC 472
 OY 245 SerAsnArgGluSerArgIleLeuPheSerThrTrpPasnLeuAspHisIleIleGluLys 264
 DB 473 ACTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 532
 OY 265 LysArgThrIleIleProThrLeuValGluAlaIleLysGluGlnAspGlyArgGluVal 284
 DB 533 AAAGCTAGAGTGTCTCCCAAGCTGCGCAGAGCTGTCAAAACAGAGAGGAGGAGGAGGAG 592
 OY 285 AspTrpGluTyrPheTyrGlyLeuLeuPheThrSerGluAsnLysLeuValHisIle 304
 DB 593 AACTGGAGTACTTTATCAGCTATGTTTACAAATGAATTAATCTAAACCTGTACATAT 652
 OY 305 ValCysHisLysLysThrHisLysLeuAsn 315
 DB 653 GCTTGCACAGAAACCAATCAATCAATCTCAGC 686
 RESULT 11
 BG705510 716 bp mRNA linear EST 07-MAY-2001
 LOCUS BG705510
 DEFINITION 602688206r1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4820856 5',
 mRNA sequence.
 ACCESSION BG705510
 VERSION BG705510.1 GI:13979919
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 716)
 NIH-MGC http://mhc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov.
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
 CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki
 Toshitsuki and Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULN at:
 http://image.llnl.gov
 Plate: LHAM10726 row: k column: 01
 High quality sequence start: 10
 High quality sequence stop: 573.
 Location/Qualifiers
 1..716
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4820856"
 /clone_11b="NIH_MGC_95"

primer adapter of sequence [5',
GAGAGAGATCTCTCGAGTATTAATATCCCCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pluescript KS(+) after bulk excision from Lambda
FLC 1. Cloning sites, 5' end: SalI; 3' end: BamHI."

BASE COUNT	123 a	183 c	186 g	122 t
ORIGIN				

ORIGIN

Alignment Scores:

Pred. No.:	2,02e-59	614
Score:	585.00	Matches: 112
Percent Similarity:	84.52%	Conservative: 19
Best Local Similarity:	72.26%	Mismatches: 24
Query Match:	32.70%	Indels: 0
DB:	10	Gaps: 0

US-09-748-451-2 (1-338) x BB619078 (1-614)

[illegible]

RESULT 13
8138
744 bp
mRNA linear
EST 19-APR-2002
A2448138
A2448138 riken1 gallus gallus cDNA clone 18n0r1, mRNA sequence.
A2448138
A2448138

VERSION AJ448138.1 GI:20215359

SOURCE	chicken.
ORGANISM	Gallus gallus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

REFERENCE 1 (bases 1 to 744)
AUTHORS Buerstedde, J.M.

COMMENT Contact: Buerstedde JM

Cellular Immunology
Heinrich-Pette-Institute

Martinistr. 52, 20251 Hamburg, Germany
Email: <http://genetics.hpi.uni-hamburg.de/dt40est.html>

FEATURES	Location/Qualifiers
source	1. .744

Source

1. : 744

1. : 744

```

/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="18n6r1"
/clone_1lb="r1ken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
note="CB inbred strain"

```

BASE COUNT	145 a	191 c	275 g	129 t	4 others
ORIGIN					

ORIGIN

Alignment Scores:

Pred. No.:	744	Length:	576
Score:	576.50	Matches:	1220
Percent Similarity:	70.00%	Conservative:	39
Best Local Similarity:	53.04%	Mismatches:	65
Query Match:	32.22%	Indels:	4
DB:	9	Gaps:	3

US-09-748-451-2 (1-338) x AJ448138 (1-744)

QY	6	lysSerValIleuArgIleValLeuArgSerProArgIysPheGlyValAlaGlyIArgSer	25
Db	40	AGGGGGTCCGCTCCGGCCGCCCGGACGCGGCGAGAGATTCCGGCGCGCGCGGAGC	99
QY	26	CysGlnGluValIleuArgIysGlyCysLeuArgPheGlnLeuProGlnAlaGlySerArg	45
Db	100	CTCCGCGGGCTCTCGCGAAGGCGTCCGCGCTGCGAGCTCCGCTTGCGAGCGCCG	159
QY	46	LeuCysLeuTyrGlnuSpGlyThrGluIleuThrGluAspTyrPheProSerValProAsp	65
Db	160	CTTGCGCTTACGAGAGCGGAGCGAGAGCTGACGAGCGAGCGCTTCTCCGACCTCCGCG	219
QY	66	AsnAlaGluValIleuValIleuThrLeuGlyGlnAlaTyrGlnGlyTyrValSerAspIle	85
Db	220	CAGACGAGACTGCTGCTCGCGCGCCGCGGAGAGACTGCGCGCTCTCGCGGAGACTG	279
QY	86	ArgArgPheLeuSerLysPheHisGluProGlnValGlyLeuIleGlnAlaGlnGln	105
Db	280	GAGCGCTCTCGCGCTGCGCTGCGACCGCGGAGCGCGGCTGTGTGGAGCGCGCGCGG	339
QY	106	IleuLeuCysAspGluGlnAlaProGlnArgGlnArgLeuAlaAspLeuIleuAlaAsn	125
Db	340	CTGCTGAGAGAGAGCGCGCGCGCGGAGCGAGAGGCTGTGCGGATCTATCCAAAC	399
QY	126	ValSerGlnAsnIleAlaGluThrArgAlaGluAspProTyrProThrPheGlnGlyLeu	145
Db	400	CTGAGCGGACACGCGCGCGGAGCGCGGAGCGAGCGAGGAACTGGTTCGAGGGCTCC	455
QY	146	GluSerArgPheGlnIleuSerLysSerGlyTyrLeuAlaGlyTyrSerCysGluSerArgIleArg	165
Db	460	GAACTCGTTTTAAGACAAATCGAGCTATATGCGGTACACTGTGAAACGAGAAATGCA	519
QY	166	SerTyrLeuArgGluValSerSerTyrTyrProSerThrValGlyAlaGluAlaGlnGlu	185
Db	520	AGCTAAATAAAGAGTGTAGTAAATTTTAAATTTTCAATCTCTACGACGAGGAAGCA	579
QY	186	PheLeuArgValIleuGly-SerMetCysGlnArgLeuArgSerMetGlnTyrAsnGlySer	205
Db	580	TATTAAGAAGATCTTACCTTATGCTGATGTGCGATTAATCTGAAATCTGAAATGTAACGGCAG	639
QY	205	rTyrPheAspArg--GlyAlaIysGlyGlySerArgLeuCysThrProGlnGly--Tr	224
Db	640	CTACTTTGACAGACGACAGAGAGGAGCGCGGCGATCGGCTNTTCAAAAGAGGAGGAATGC	699
QY	223	pPheSerCysGlnGlyPro--PheAsp	231
Db	700	TTTTTTTGTGCAAGGGCCCTTTTGTAC	727

LOCUS	DEFINITION	AL602837	637 bp	MRNA	linear	EST 14-ND
LOCUS		AL602837	637 bp	MRNA	linear	EST 14-ND
DEFINITION		DKFZP660116.r1	686 (synonym: h1cc3)	<u>Homo sapiens</u>	<u>CDNA clone</u>	
		DKFZP660116.5,	mRNA sequence.			

Db 293 TATGTGATGACATCGGCTCTCCATGCTTTTCGC----- 255
 QY 101 GlnAlaIaIaGlnGlnLeuLeuCyAspGluGlnAla-ProGlnArGlnArGlnLeuA1 120
 Db 254 CAGCC-GCCGGGACGCTACTGTGTGATGACAGGCCGCCACAGAGGACGACGCGGC 196
 QY 120 aAspLeuLeuH1aSnValSerGlnaSnIleAlaIaGlnThraGlnaGlnaAspPro 140
 Db 195 CGACCTCTCTACACCTCAGACCCAGACGTCGACGACCCGCGCTGAGACCCGCT 136
 QY 140 CTTPheGluGlnLeuLeuSerArgPheGlnSerGlyGlyLeuArGlyrGlyrSercy 160
 Db 135 GTGCTTGAAGGCTGAGATCCGATTTTGATTAAGCTGCTGCTATCTGAGATACGCTG 76
 QY 160 sGluSerArgIleArGSerTyLeuArGlnuValSerSerTyLeuProSerThraValGlyA1 180
 Db 75 TGAAGCCGCGATCCGAGGTTACCTGACAGAGGATGAGTCCGTCCTCCATGCGGGTGC 16
 QY 180 agluAlaGlnGlu 184
 Db 15 AGAGGCTCAGAG 3
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 ORGANISM
 SOURCE
 KEYWORDS
 EST 26-JAN-2001
 515 bp mRNA linear
 H3124D06-5 NIA Mouse 15K CDNA Clone Set Mus musculus CDNA clone
 H3124D06-5', mRNA sequence.
 B0086312
 B0086312
 B0086312.1 GI:12568876
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 515)
 Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jarradt, S.A., Tanaka
 J.T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of NIA mouse 15K CDNA clone set
 unpublished (2001)
 Other ESTs: H3124D06-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@nigun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please
 visit <http://lgsun.grc.nia.nih.gov/CDNA/15k.html> for details.
 Plate: H3124 row: D column: 06
 Seq primer: -21M13 Reverse
 High quality sequence stop: 515
 POLYA-No.
 Location/Qualifiers
 1. 515
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="H3124D06"
 /sex="Clones arrayed from a variety of CDNA libraries"
 /dev_stage="Clones arrayed from a variety of CDNA
 libraries"
 /lab_host="DH10B"
 /note="Vector: pSPORT1, Site 1: SalI, Site 2: NotI; This
 clone is among a rearrayed set of 15,247 clones from 11
 embryo CDNA libraries (including preimplantation stage
 embryos from unfertilized egg to blastocyst, embryonic
 part of E7.5 embryos, extraembryonic part of E7.5 embryos
 and E12.5 female mesonephros/gonad) and one newborn
 ovary CDNA library. Average insert size 1.5 kb. All
 source libraries are cloned unidirectionally with Oligo(dT
)-Not primers. References include: (1) Genome-wide
 expression profiling of mid-gestation placenta and embryo

using a 15,000 mouse developmental CDNA microarray, 2000,
 Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
 large-scale CDNA analysis reveals phased gene expression
 patterns during preimplantation mouse development, 2000,
 Development, 127: 1737-1749; (3) Genome-wide mapping of
 unselected transcripts from extraembryonic tissue of
 7.5-day mouse embryos reveals enrichment in the t-complex
 and underrepresentation on the X chromosome, 1998, Hum
 Mol Genet 7: 1967-1978.
 BASE COUNT 102 a 163 c 148 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,84e-55 Length: 515
 Score: 551.00 Matches: 104
 Percent Similarity: 85.31% Conservative: 18
 Best local Similarity: 72.73% Mismatches: 21
 Query Match: 30.80% Indels: 0
 DB: 12 Gaps: 0
 US-09-748-451-2 (1-338) x B0086312 (1-515)
 QY 1 MetLeuGlnYsProlYsSerValLysLeuArGlnArGSerProArGlySphGly 20
 Db 71 GTGCTCCGCCAACCCTAATGCGTCAAGTTCGAGCCCTACATACGCCCTCGCAAGTTCGGC 130
 QY 21 ValAlaGlnArGSerCySglnGlnuValLeuArGlySglnCySLeuArPheGlnLeuPro 40
 Db 131 GTGGCGGCGCCGAGACTGCGAGACCTGCTGCGTAAAGGCGTCCGTCCTCCACTCCG 190
 QY 41 GlnArGlySerArGLeuCySLeuTyrgLnuAspGlyThrGluThrGlnuAspTyRpe 60
 Db 191 ATGCCCGGTTCCCGCTGCTGCTTACGAGATGCGACGAGCGAGCGAGCGAGTCTTC 250
 QY 61 ProSerValProAspAsnAlaGlnuValLeuArGlnuLeuThraGlnuAlaTrpGlnGly 80
 Db 251 CCGGCGCTTCCCAACGACGAGCTGACCTCTATGCTACCCGCTGCGAGACTGCGATGCG 310
 QY 81 TyrValSerAspIleArGArGpHeLeuSerAlaPheHsGlnuProGlnuValGlyLeuile 100
 Db 311 TATGTGATGATGACATCAGACGCTTCTCAGTGTGTTATAGCCACATGCGCGGTGATC 370
 QY 101 GlnAlaIaIaGlnGlnLeuLeuCyAspGluGlnAla-ProGlnArGlnArGlnLeuA1 120
 Db 371 CAGGCTCAGCGGCAACGCTGTCAGATGAGAGGCCGCCACTGACGCAAAAGCTGCTGCC 430
 QY 121 AspLeuLeuH1aSnValSerGlnaSnIleAlaIaGlnThraGlnaGlnaAspPro 140
 Db 431 GATCTTGTGATCAGCTGAGGACGAGATATTTACTGACAGACCGCGGAGACCCATCC 490
 QY 141 TrpPheGlu 143
 Db 491 TGGTTGAA 499
 RESULT 17
 B1789500
 LOCUS
 DEFINITION
 1695e08.y1 Melton Normalized Mixed Mouse Pancreas 1 NT MMS1 MMS
 musculus CDNA clone IMAGE:5661422 5' similar to TR:054788 054788
 CAD PROTEIN. ; mRNA sequence.
 B1789500
 B1789500.1 GI:15817225
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 537)
 Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Scarce, M., Brestelli, J., Girdwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, D., Blisstein, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
 , M., Gibbons, M., McCann, R., Cole, R., Tsagarelashvili, R., Williams, T.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	Location/Qualifiers
Source	1. .643

BASE COUNT	133 a	183 c	195 g.	132 t
ORIGIN				

Alignment Scores:

Pred. No.:	2.07e-51	Length:	64
Score:	518.50	Matches:	10
Percent Similarity:	80.13%	Conservative:	10
Best local Similarity:	67.31%	Mismatches:	29
Query Match:	28.98%	Indels:	2
DB:	10	Gaps:	1

US-09-748-451-2 (1-338) x BB628807 (1-643)

[illegible]

Db 499 TCCTGGGTTGGAAGATCCCCCACTTTTAAACCAAGGGGGG 544

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
19	BE755721	513 bp	2BOV	CDNA	linear	EST 25-APR-2001	Bos taurus
	BE755721		2BOV	CDNA	linear	EST 25-APR-2001	Bos taurus
	BE755721		2BOV	CDNA	linear	EST 25-APR-2001	Bos taurus
	BE755721.1		2BOV	CDNA	linear	EST 25-APR-2001	Bos taurus
	EST.						Bos taurus
	COW.						Bos taurus

BASE COUNT	90 a	152 c	177 g	94 t
ORIGIN				

Alignment Scores:

Pred. No.:	3,81e-48	Length:	516
Score:	490.00	Matches:	96
Percent Similarity:	83.85%	Conservative:	13
Best local Similarity:	73.85%	Mismatches:	21
Query Match:	27.39%	Indels:	0
DB:	12	Gaps:	0

US-09-748-451-2 (1-338) x BE755721 (1-513)

Qy	1	MeLeuGlnIysProIysSerValIysLeuAlaGalaLeuArgSerProArgIysPheIy	20
	1122	GGCTCCGGAGGCCAAGAGCTTCAAGAGCTGGCGTCCGTGCATAGCGAGAGAA	18
Db	21	ValAlaGlyAArgSerCysGlnGluValLeuAArgIysIyCysLeuArgPheGlnLeu	40
Qy	182	GTCGGCGGACGAGAGCTGCGAGAGAGTGGCTGCGGAAAGGGGTGCACACG	24
Db	41	GluArgGlySerArgLeuCysLeuIytrGluAspGlyThrGluLeuThrGluAsp	60
Qy	242	ATCCCTGGATCCCGCCGTGTCTCTATATGAGATGACACGGAATTATCGAGAT	30

```

OY 61 ProserValProAspAsnAlaGluLeuValLeuThrLeuGlyAlaAlaTrpGlnGly 80
DB 302 TGAAGTCCCCGACACATCATGCTGTGCTTCTCTACCTCCGCGACAGCTGGAGAGG 361
OY 81 TyrValSerAspIleArgRphLeuSerAlaPheHisGluProGlnValAlaGlyLeu 100
DB 362 TTCGTAGTACATCATGCGCTTCTCTGAGCGTGTCCAGAGCCGCGATCCAGTCTCATC 421
OY 101 GlnAlaAlaGlnGluLeuLeuGlyAspGluAlaAlaProGlnAlaArgGlnLeuAla 120
DB 422 CAGCGCCCGCCGCGAGCTGTCTGCGAGCGAGCGGCGCCACTGCGGAGAGAGCTGCGCC 481
OY 121 AspLeuLeuHisAsnValSerGlnAsnIle 130
DB 482 GACCTGCTGGGCGACCGTCAATGAGAACATC 511

LOCUS AT 20
DEFINITION B026619 688 bp mRNA linear EST 05-DEC-2001
ACCESSION B026619 MFOISSA CDNA Oryzias latipes CDNA clone MFOISSA174E05 3',
VERSION B026619
KEYWORDS B026619.1 GI:17391478
SOURCE EST.
ORGANISM Japanese medaka.
Oryzias latipes.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
1 (bases 1 to 688)
REFERENCE Kohara, Y., Shin-I, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.
TITLE Medaka EST project in Takeda's lab
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
FEATURES
SOURCE
1.688
/organism="Oryzias latipes"
/strain="Hd-tr"
/db_xref="taxon:8090"
/clone="MFOISSA174E05"
/clone_lib="MFOISSA CDNA"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"

BASE COUNT 133 a 172 c 178 g 205 t
ORIGIN

Alignment Scores:
Pred. No.: 1.21e-47 Length: 688
Score: 487.50 Matches: 90
Percent Similarity: 71.60% Conservative: 26
Best Local Similarity: 55.56% Mismatches: 43
Query Match: 27.25% Indels: 3
DB: 13 Gaps: 1

US-09-748-451-2 (1-338) x B026619 (1-688)
OY 180 AlaGluAlaGlnGluLeuValLeuThrLeuGlySerMetCysGlnArgLeuArgSer 199
DB 668 GCCAAGTCCAGAGAGATATTCAGAGCCCTTCAAGCTCTGTAGAAATGCTGAAGGT 609
OY 200 MetGlnTyrAsnGlySerTyrPheAspArgGlyAlaAlaGlySerArgLeuGlyThr 219
DB 608 GACGGTTATACGGCGAGCTACTTCACAGAGAGCCAGAGAACGGCGCGCTCTGACAG 549

```

```

OY 220 ProGlnGlyTyrPheSerCysGlnGlyProPheAspMetAspSerCysLeuSerArgHis 239
DB 548 CAGAGGAGATGGTTACACCTGCGCAGAGGACCGTTTACCAAGAGAGCGCCAGAGCTCAC 489
OY 240 SerIleAsnProTyrSerAsnArgGluSerArgIleLeuPheSerThrTrpAsnLeuAsp 259
DB 488 TCCATCAACCCCTACAGCAGCAGGAGAGCGAGATCATCTTCAGACCTGGAACCTGGAC 429
OY 260 HisIleLeuGluTyrAsnArgThrIleLeuProThrLeuValGluAlaIleLeuGln 279
DB 428 CACAGATTCAGAAAGAGAGAGCGTCATTCACAGCGTGTGGAGCTCTTAAGACTCAC 369
OY 280 AspGlyArgGluValAspTrpGlnTyrPheTyrGlyLeuLeuPheThrSerGlnAsnLeu 299
DB 368 CAGAGCTCAGCGGTCAACCTGAGACTATTTCTACGGCTGCTGTTCACAGAGAACTTG 309
OY 300 LysLeuValHisIleValCysHisLysLysThrThrHisLysLeuAsnCysAspProSer 319
DB 308 AAGTTGGTTCACATCGTTGTCCAGAAAGAGAGTCACAACTTCATGCGACCGCCAGC 249
OY 320 ArgIleTyrLysProGlnThrArg-----LeuLysArgLysGlnProValArgLys 336
DB 248 ATGGCGGTGAGGGCTCCAGCCAGCGAGGGGAGTGAAGAAAGGTCCAGGTGAAGAG 189
OY 337 ArgGln 338
DB 188 AGCGCA 183

```

Search completed: May 26, 2003, 16:52:57
 Job time : 1715 secs

THIS PAGE BLANK